Aalto University

School of Electrical Engineering

Space Science and Technology

Master's thesis – Diplomityö - Diplomarbete

Identification of narrow-line Seyfert 1 galaxies in a spurious source sample using spectral line profiles

Irene Björklund

Thesis supervisor and advisor Professor Anne Lähteenmäki

Aalto University
School of Electrical Engineering
Master's Program in Nano and Radio Sciences

Abstract of the
Master's Thesis
in English

Author Irene Björklund

Title of thesis Identification of narrow-line Seyfert 1 galaxies in a spurious source sample using spectral line profiles

Program Space Science and Technology **Code of program** ELEC3039

Supervisor and advisor Anne Lähteenmäki

Date 07.05.2019 **Language** English **Number of pages** 8 + 77

Abstract in English

This thesis is a compilation of both general knowledge on galaxies and active galaxies as well as detailed knowledge on NLS1 galaxies. A recent research paper claimed that they confirmed 11001 sources to be NLS1 galaxies. The main object of this thesis was to determine how many of these sources are truly NLS1 with a high certainty.

The sources were deemed most likely NLS1 galaxies by studying the emission lines of the galaxy spectra and performing mathematical analysis as well as observing the spectra by eye. After the sources had been either deemed most likely NLS1 or discarded, only 3998 sources remained.

The most likely NLS1 galaxy sources were divided into two categories based on their profile type: Lorentzian (Cauchy) or Gaussian. 11 parameters of the most likely NLS1 galaxy sources were further analyzed to find possible differences between the two profile types. With the results of this thesis, researchers can study NLS1 galaxies better by not having a contaminated source sample and can hopefully find instrumental properties of NLS1 galaxies.

Key words NLS1 galaxy, galaxy spectra, emission lines, active galaxies

Aalto University
School of Electrical Engineering
Master's Program in Nano and Radio Sciences

Abstract of the
Master's Thesis
in Swedish

Författare Irene Björklund

Titel Identifiering av typ 1 Seyfertgalax med smala spektrallinjer från ett falskt källurval genom att använda spektrallinjeprofiler

Huvudämne Rymdfysik och teknik **Huvudämnets kod** ELEC3039

Handledare och ansvarig lärare Anne Lähteenmäki

Datum 07.05.2019 **Språk** Engelska **Sidoantal** 8 + 77

Abstrakt på Svenska

Detta diplomarbete är en samanställning av både generell information om galaxer och aktiva galaxer samt av mer detaljerad information om NLS1 galaxer. En färsk forskning påstå att de bekräftade 11001 källor att vara NLS1 galaxer. Huvudsyftet med detta diplomarbete är att avgöra hur många av dessa källor är verkligen NLS1 galaxer med hög säkerhet.

Källorna ansågs vara högst sannolikt NLS1 galaxer genom att studera emissionslinjerna av källornas galaxspektra och genom att utföra matematisk analys samt genom en visuell analys av spektra. Efter att källorna var antigen betraktade NLS1 eller förkastade, fanns det bara 3998 källor kvar.

De högst sannolikt NLS1 galaxkällor delades till två kategorier baserade på deras profiltyp: Lorentz (Cauchy) eller Gaussisk. 11 storheter av de högst sannolikt NLS1 galaxkällor var ytterligare analyserade för att hitta möjliga olikheter mellan de två profiltyper. Med resultaten från detta diplomarbete kan forskare studera NLS1 galaxer bättre utan att ha ett kontaminerat källurval. Forskarna kan då hoppeligen hitta väsentliga egenskaper för NLS1 galaxer.

Nyckelord NLS1, galaxspektra, emissionslinjer, aktiva galaxer

Acknowledgements

I would like to thank not only my supervisor Anne Lähteenmäki for providing useful feedback and tips for my thesis as well as giving me this topic in the first place, but also Marco Berton for providing loads of useful information about NLS1 and spectral lines of active galaxies. I would also like to thank Marco for providing me with the code that I used for obtaining values for the different sources and helping me to see if the values I obtained made any sense and if my further calculations on these values seemed reasonable. I would also like to thank Emilia Järvelä for her valuable input on this topic. Last but not least, I want to thank Sean Connor for dealing with me banging my head to the wall when I couldn't figure something out.

Contents

Abstract in English i

Abstract in Swedish ii

Acknowledgements iii

Contents iv

Symbols, constants, concepts, and abbreviations vi

Symbols and constants vi

Concepts and abbreviations vii

Introduction 1

1. Astronomical units 3

1.1 Distance 3

1.2 Mass 4

1.3 Time and speed 4

2. What are galaxies? 6

2.1 Features and formation 6

2.1.1 Redshift 7

2.2 Classification 8

2.2.1 Elliptical 9

2.2.2 Lenticular 9

2.2.3 Spiral 10

2.2.4 Irregular 10

2.2.5 Hubble's galaxy classification map 11

2.2.6 Other type of classification 12

2.3 Galaxy evolution 12

3. Active galaxies 15
3.1 Jets of galaxies 17
3.2 Active galaxy classification 18
3.3 Spectral lines of active galaxies 18
3.4 Forbidden lines 20
4 Unified model of active galaxies 21
5 Radio galaxies 24
6 Quasars 27
7 Blazars 28
8 Seyfert galaxies 29
8.1 Type 1 Seyfert 31
8.2 Type 2 Seyfert 33
9 Narrow-line Seyfert-1 galaxies 35
9.1 Spectrum of a NLS1 36
9.1.1 Broad-line Seyfert-1 38
10 The sample and method of analysis 39
11 Results 46
12 Discussion 60
13 Conclusions 62
References 64
Appendix – Marco Berton's source analysis code 68
Appendix – Marco Berton's luminosity code 73

Symbols, constants, concepts, and abbreviations

Symbols and constants

m	meter
AU	Astronomical Unit, $1\ \mathrm{AU}\ =\ 149\ 597\ 870$ meters
ly	light-year, $1\ \mathrm{ly}\ =\ 63\ 241$ au
pc	parsec, $1\ \mathrm{pc}\ =\ 3.26\ \mathrm{ly}$
kg	kilogram
$\mathrm{M}_\odot$	Solar mass, $\mathrm{M}_\odot = 1.989 * 10^{30}$ kg, $\mathrm{M}_\odot = 332\ 950\ M_\oplus$
nm	nanometer, $1\ \mathrm{nm} =\ 10^{-9}\ \mathrm{m}$
Å	Ångström, $1\ \text{Å} = 0.1\ \mathrm{nm}$
pm	picometer, $1\ \mathrm{pm} =\ 10^{-12}\ \mathrm{m}$
µm	micrometer, $1\ \mu\mathrm{m} =\ 10^{-6}\ \mathrm{m}$
mm	millimeter, $1\ \mathrm{mm} =\ 10^{-3}\ \mathrm{m}$
L	luminosity
$\mathrm{L}_\odot$	luminosity of the Sun, $\mathrm{L}_\odot = 3.828 * 10^{26}$ W
c	speed of light, $c =\ 299\ 792\ 458\ \mathrm{m/s}$
k	kilo, $\mathrm{k} = 10^3$
z	redshift
γ	gamma
α	alpha
β	beta
G	Newton's gravitational constant, $G = 6.67 * 10^{-11}\ \frac{\mathrm{Nm^2}}{\mathrm{kg^2}}$
eV	electron volt, $1\ \mathrm{eV} = 1.602 * 10^{-19}\ \mathrm{J}$
H_0	Hubble's constant, still being refined, estimated as $71\ \frac{\mathrm{km*Mpc}}{\mathrm{s}}$
M	mega, 10^6
H	hydrogen
Fe	iron
O	oxygen
M_{BH}	mass of black hole
$\mathrm{M}_{\mathrm{Edd}}$	Eddington mass limit
$\mathrm{L}_{\mathrm{Edd}}$	Eddington luminosity
χ^2	reduced chi-squared statistic

Concepts and abbreviations

AGN	Active Galactic Nuclei
X-ray	electromagnetic radiation, wavelength range: 0.01 nm and 10 nm
γ-ray	electromagnetic radiation, wavelength is 10 pm or less
cD	giant elliptical galaxy class
E0, …, E7	elliptical galaxy classes
S0	lenticular galaxy class
Sa-Sb-Sc	normal spiral galaxy classes
SBa-SBb-SBc	barred spiral galaxy classes
Irr I	irregular galaxy 1
Irr II	irregular galaxy 2
dIrr	dwarf irregular galaxy
dSph	dwarf spheroidals
dE	dwarf ellipsoids
uFd	ultrafaint dwarfs
BCD	blue compact dwarfs
UCD	ultracompact dwarfs
NLS1	narrow-line Seyfert-1 galaxy
BLS1	broad-line Seyfert-1 galaxy
IR	infrared
UV	ultraviolet
QSO	quasi-stellar object
Light-day	one 365.25th of a light-year $2.59 * 10^{13}$ m
BLR	BLR
NLR	NLR
Protogalaxy	a gas cloud that is forming into a galaxy
Bulge	a dense, tightly packed, swarm of stars
Superluminal	faster-than-light
Hard X-ray	highly energized X-ray radiation, energy over 10 keV
Soft X-ray	lowly energized X-ray radiation, energy less than 10 keV
FIR	far infrared radiation, the wavelength is between 15 μm and 1 mm
Nebula	interstellar cloud of dust, H, He, and other ionized gases. Has been used as a term for any diffuse astronomical object
Forbidden lines	spectral lines that occur due to electron transitions that are

	not generally permitted by the quantum mechanics selection rules
Synchrotron radiation	nonthermal radiation generated by charged particles, that are spiraling around magnetic field lines at speed close to the speed of light
Accretion disk	disk-like flow of particles, gas, dust, or plasma around an astronomical object
Redshift	phenomenon that occurs when the wavelength of electromagnetic radiation of an object increases
Blueshift	opposite of redshift, wavelength decreases
Torus	doughnut shaped gas cloud rotating near the galaxy core
Lyman series	hydrogen spectral series, electrons move from an excited state to the lowest energy state. The wavelengths are always in the ultraviolet band
Balmer series	hydrogen spectral series, electrons move from an excited state to the second lowest energy state
BL Lac	BL Lacertae, a subclass of blazars
OVV	optically violently variable
FWHM	full width at half maximum
H_β	hydrogen Balmer series
Fe II	singly ionized iron
[O III]	doubly ionized oxygen
MBH	massive black hole
BBB	big blue bump
Eddington luminosity	maximum luminosity of a celestial body
Eddington mass limit	a (theoretical) mass upper limit of a celestial body
SDSS	Sloan Digital Sky Survey
DR	Data Release
S/N	Signal-to-Noise
FSRQ	flat-spectrum radio quasar

Introduction

"Man must rise above the Earth—to the top of the atmosphere and beyond—for only thus will he fully understand the world in which he lives."
- Socrates

The ever-expanding Universe is a concept that scientists are keen to understand. Throughout the history of mankind, humans have learned a great deal of our own planet, but much of everything that is beyond Earth is still a mystery. Though many have already looked at the skies and performed calculations beyond belief, with at times no instruments to aid, there is still a lot to learn. Today scientists study the Universe, seeking knowledge of many different sorts. The knowledge that I seek in this thesis is on narrow-line Seyfert 1 (NLS1) galaxies. There are relatively few galaxies that can be said to be NLS1 galaxies with a high confidence level. There are three goals that I achieve in this thesis. The first goal of this thesis is to determine how many of the galaxy sources are NLS1 galaxies from a spurious dataset. The data used in this thesis is obtained from Sloan Digital Sky Survey (SDSS) and is the same data that Rakshit et al. published in 2017. This data is used because Rakshit et al. stated that all the galaxy sources in the sample are NLS1 galaxies, but it was noticed that this is not the case as the spectra from many sources was not the spectra of a NLS1. Determining which sources are most likely NLS1 galaxies is vital, as it is impossible to otherwise properly tell the real properties of NLS1 galaxies. With the most likely NLS1 sources, it is possible to start determining fundamental properties of this galaxy type.

To reach the goals of this thesis, 11001 sources are analyzed. Of these 11001 sources, the most likely NLS1 galaxies are determined by the utilization codes provided by Marco Berton and through hands-on analyzation based on the results from the code. The used code can be found in the appendix (Appendix – Marco Berton's source analysis code). Programs like Python, Excel, and MATLAB have been used in this thesis. After finding which sources are most likely NLS1 galaxies, the sources are divided into two categories based on their fit: Gaussian or Lorentzian. Dividing the sources into two categories is the second goal of this thesis. The final goal of this thesis is to analyze integral parameters of the most likely NLS1 galaxies. A total of 11 parameters are studied in this thesis of the most likely NLS1 sources. The considered parameters are: Redshift, black hole mass, R4570 value, Hβ flux, Hβ luminosity, [O III] flux, [O III] luminosity, Eddington ratio, monochromatic 5100 Å luminosity, bolometric luminosity, and FWHM Hβ. The parameters are obtained by the utilization of equations from the article by Rakshit et al. The studied parameters have been chosen because they provide information on the properties of the sources.

This is a master's thesis for Aalto University in the field of space science and technology and to the Metsähovi Observatory. Before getting to the main topic of this thesis fundamental topics will be presented. These topics are: astronomical units, what galaxies are, different classification of galaxies, galaxy evolution, active galaxies and the different classes of them (radio, Seyfert galaxies, quasars, and blazars), unified model of active galaxies, and more detailed knowledge on NLS1 galaxies. These topics have been chosen to make it possible to understand better the main topic and object of this thesis. Results, discussion, and conclusions can be found at the end of the thesis.

Three books have been used for the general knowledge parts in this thesis. These books are *Fundamental Astronomy* by Karttunen et al. (2016), *Galaxies in the Universe* by Sparke & Gallagher (2007), and *Active Galactic Nuclei* by Robson (1996). With this mention, the general knowledge parts will not have reference markings from these books.

1. Astronomical units

Distance, mass, time, and speed are all important measurements needed both on Earth and when researching what lies beyond Earth. On Earth these physical quantities are measured in meters, kilograms, seconds, and meters per second. These units more commonly known as the SI units. The SI units are easy to understand and remain relatively easy to comprehend on Earth. When researching and discussing objects and phenomena in space, the SI units can be applied, but they become hard to fathom. The distance between Earth and the Sun is roughly 149 597 870 700 meters, or roughly 150 million kilometers. This distance is very hard to fully grasp, as well as hard to remember. The Sun lies relatively close to Earth, which can be hard to be believe when thinking about the distance in SI units. Many stars and galaxies lie much further away from Earth than the Sun. The furthest space object humans can perceive from Earth, with the naked eye, is the Andromeda galaxy. The Andromeda galaxy is located $2.4 * 10^{19}$ kilometers away from Earth. Many stars and galaxies are not visible to the naked eye from Earth, which means that they are even further away from Earth, or very faint. To make it easier to understand distance and other measurements in space, new system of units is needed, the astronomical system of units, or the IAU (1976) System of Astronomical Constants. Table 1.1 has a summary on the conversion between the units.

1.1 Distance

The IAU (1976) System of Astronomical Constants includes units like the Astronomical Unit, Solar mass, and days. These units can be converted to SI units if wanted, but the astronomical system of units is more convenient in astronomy. The regular SI units can be used at times, when for example measuring the distance from Earth to a satellite in a geocentric orbit. The Astronomical Unit is a unit of length used for measuring distance between objects in space. 1 AU corresponds to 149 597 870 meters, in other words, the Astronomical Unit is the rough distance between Earth and the Sun. The Astronomical Unit is generally used for measuring distances within the Solar System. The Astronomical Unit is not well applicable when measuring very large distances.

Another very well-known unit for distance is a light-year. A light-year is the distance light travels in vacuum in one Julian year (365.25 days). 1ly is equal to $9.46 * 10^{15}$ meters or 63 241 AU. Distance can also be measured in light-days. A light-day is 365.25^{th} of a

light-year or $2.59 * 10^{13}$ m Light-years can also be broken up more, i.e. light-seconds, light-hours. Another common unit of distance in astronomy is parsec. One parsec is roughly 3.26 ly. Parsecs can also be used with the common SI unit prefixes, like kilo. The Andromeda galaxy is located 770 kiloparsecs from Earth.

Parsecs are usually only used for measuring large distances outside our Solar System and it is defined as the distance at which one Astronomical Unit subtends an angle of one second of an arc, so an arcsecond. An arcsecond (symbol “) is equal to $\frac{1}{3600}$°. As a conclusion, the usual trio of units for measuring distance in astronomy is Astronomical Unit (AU), light-year (ly), and parsec (pc). Kilometers can also be used for measuring shorter distances, like the distance between Earth and the International Space Station (ISS).

This thesis focuses heavily on spectral lines, so also small-scale distance units are needed. The common nanometers (nm) and micrometers (µm) are often used when researching spectral lines. The not as well-known unit when researching spectral lines is Ångströms. Ångströms, named after the Swedish physicist Anders Jonas Ångström, is one billionth of a meter, 0.1 nm.

1.2 Mass

The astronomical distance units are maybe the most common of the astronomical system of units, but other units are important as well, like mass. In astronomy, there are two basic units of mass, where the more common one is Solar mass. Solar mass is a very straight forward term, as it means, the mass of the Sun. The mass of the Sun is $1.989 * 10^{30}$ kg. The symbol for Solar mass is $M_{\odot}$. The $\odot$ is commonly used when signifying the Sun. For example, the symbol for Solar luminosity is $L_{\odot}$ ($L_{\odot} = 3.828 * 10^{26}$ W).

1.3 Time and speed

The SI unit for time on Earth is a second, but in the astronomical system of units, the preferred unit is day. One day is equal to 86 400 seconds. Day is preferred over seconds because in astronomy things usually do not proceed so quickly due to the large

distances. It is also possible to measure time in years, though then it is important to have the correct definition of a year. According to the International Astronomical Union, one year corresponds to one Julian year, so 365.25 days. The symbol for a year is a, which comes from the Latin word *annus*, which means a year.

Speed is perhaps one of the most similar units between the astronomical system of units and the SI units. The SI unit for speed is m/s. In astronomy, km/s is often the preferred unit due to the high velocities. The constant for the speed of light, denoted as *c*, roughly equaling 299 792 458 m/s, is a commonly used constant in astronomy.

Table 1.1 Conversion between astronomical units and SI units

Astronomical Unit	**SI unit**
1 AU	$149\ 597\ 870$ m
1 ly	$9.4605284 * 10^{15}$ m
1 pc	$3.08567758 * 10^{16}$ m
1 Å	10^{-10} m
$\mathbf{1}\ \boldsymbol{M_{\odot}}$	$1.989 * 10^{30}$ kg
$\mathbf{1}\ \boldsymbol{M_{\oplus}}$	$5.97 * 10^{24}$ kg
$\mathbf{1}\ \boldsymbol{L_{\odot}}$	$3.828 * 10^{26}$ W

2. What are galaxies?

A galaxy is a gravitationally bound system, consisting of stars, dust, dark matter, and interstellar gas. There are many different types of galaxies. Galaxies can have a million stars, but galaxies can even contain 10^{12} stars. Stars are the major light source in galaxies. Galaxies are what make the Universe. Galaxies can be viewed as the base LEGO pieces that create the entirety that we call the Universe. Galaxies can either have a simpler or a more complex construction and build.

Galaxies can be nearly solely made of neutral gas, but galaxies can also consist of only stars, with no distinct individual features. There are galaxies that consist of multiple separately interacting components. These components can be stars, dust, cosmic rays, molecular clouds, magnetic fields, as well as neutral and ionized gas. Galaxies can be in small groups or create large clusters. There are galaxies that have a core that emits so much radiation that it can inundate the normal radiation of the galaxy. These galaxies are known as active galaxies.

The general properties of active galaxies will be discussed in Chapter 3. Chapter 3 will also explain the spectrum of an active galaxy, as well as about relativistic jets. Three active galaxy types are discussed in their respective chapters, Chapter 5 is about radio galaxies, Chapter 6 is about quasars, and Chapter 7 is about blazars. Chapter 4 discusses the unification model of active galaxies. Chapter 8 and 9 focus heavily on Seyfert galaxies, the most important type of active galaxies for this thesis.

2.1 Features and formation

The confirmation on the existence of dark matter has been a helpful factor on creating models for both the origin of galaxies as well as the evolution of galaxies. The confirmation of dark matter also changed how researchers viewed galaxies as well as the definition of galaxies. Researchers today consider a galaxy to be a star system if and only if it also includes dark matter on top of stars. This definition also helps to determine when observing very small galaxies that could be either star clusters or galaxies.

The general structure of galaxies varies with the morphology of a galaxy, but most galaxies have a central supermassive black hole. A large portion of disk galaxies have a

central bar. The central bar, located inside many disk galaxies, causes gravitational instabilities and can alter the galaxy's evolution. All galaxies that have a prominent disk are known as disk galaxies. Spiral and lenticular galaxies have prominent disks and are thus disk galaxies.

The luminosity of galaxies varies greatly. Usually, the luminosity of galaxies or their cores are compared to the Solar luminosity. The brightest non-active galaxies can be as luminous as $10^{12}\mathrm{L}_{\odot}$. The faintest of the known galaxies have a luminosity of $10^{5}\mathrm{L}_{\odot}$. The luminosity of most non-active galaxies is less than $10^{12}\mathrm{L}_{\odot}$.

Another commonly varying feature of galaxies is the mass of the galaxy. Like with the luminosity of galaxies, also the mass of galaxies is compared to the Solar mass. Determining the mass of the galaxy can be tricky, as galaxies do not have very clear ends, but slowly fade out in the Universe. To avoid having to guess where the galaxy ends, only the bright central parts are usually considered. With this limiting factor, dwarf galaxies have an average mass of $10^{7}\mathrm{M}_{\odot}$ and giant galaxies have an average mass of $10^{13}\mathrm{M}_{\odot}$. The radii of galaxies face the same problem as with determining the mass. To once again avoid having to guess, the same limiting factor is placed as with the mass of galaxies. With the limiting factor in affect, the average radius of dwarf galaxies is 0.5 kpc, while the average radius of giant galaxies is 30 kpc.

2.1.1 Redshift

Redshift is a phenomenon in which the wavelength of electromagnetic radiation moves toward a longer wavelength. There are three main types of redshift: gravitational redshift, cosmological redshift, and peculiar redshift. The phenomenon is named redshift because red is the longest wavelength on the visible spectrum. The opposite effect of redshift is blueshift, in which the wavelength decreases.

Gravitational redshift occurs when electromagnetic radiation escapes a gravitational well and the wavelength increases. For the photons to able to escape, they must spend energy. As per Einstein's general relativity theory, photons cannot decrease their speed from the speed of light, the energy loss must occur through a change of frequency, and thus through the increase of wavelength.

Cosmological redshift is caused by the Universe constantly expanding. When the wavelength becomes longer, the object that radiates the wavelength moves away from

the observer. Due to redshift, the spectral lines observed of for example a galaxy, are all shifted to the longer wavelengths. Redshift can tell how far a stellar object is from us.

Peculiar redshift is related to peculiar velocity of an astronomical object to a rest frame. This type of redshift can tell the observer about the velocity of the astronomical object relative to the observer. Peculiar redshift is caused by the relativistic Doppler effect. According to Hubble's law, the larger the redshift the further away the object is from the origin point. Hubble's law can be seen in Equation 3.1.

$$v = H_0 d \tag{3.1}$$

In Equation 3.1, v is the velocity of recession measured in $\frac{\text{km}}{\text{s}}$. H_0 is the Hubble's constant measured in $\frac{\text{km}*\text{Mpc}}{\text{s}}$. The value for the Hubble's constant is still being refined, but it is generally around $71\ \frac{\text{km}*\text{Mpc}}{\text{s}}$. The last variable d stands for the distance to the galaxy measured in Mpc.

2.2 Classification

Galaxies can be classified in different ways depending on what feature we are looking at. Dwarf and giant galaxies are obvious classifications based on the size and mass of the galaxies. There are also other classification categories. One of the perhaps most common classification methods for galaxies is classification based on the morphology of galaxies. The problem with categorizing galaxies by their morphology is that it is to an extent subjective rather than objective. Nonetheless, morphology categorization is a very good basic method of classifying galaxies and provides a framework. Morphological categorization does not work for all galaxies either. The galaxies that morphological classification applies to, have to be both bright and have a large radius, so that they are visible in the night sky. If a galaxy has too large of a radius compared to its luminosity, it will not be visible on the night sky. If the galaxy has a very small radius, it will appear like a star.

The most common galaxy classification system was made by Edwin Hubble. Edwin Hubble proposed in 1926 that there are three main types of galaxies: elliptical, lenticular, and spiral. These can be further divided into sequences, for example the spiral galaxies are divided into two sequences: normal and barred spirals. On top of the three main

categories, Hubble included one more category, the irregular galaxy category. At the end of section 2.2.4, Figure 2.1 portrays the Hubble classification scheme.

2.2.1 Elliptical

Elliptical galaxies are round, smooth, and nearly featureless. The nearly featureless part means that there are no spiral arms or apparent dust lanes. Elliptical galaxies have very few young stars. When observing an elliptical galaxy in the sky, it will appear as a concentration of stars in an elliptical form. There are eight different categories for elliptical galaxies, E0, …, E7. E0 looks, though an elliptical galaxy, like a circle in the sky. The type of the elliptical galaxy is defined as

$$n = 10\left(1 - \frac{b}{a}\right) \tag{2.1}$$

Where *a* is the major axis of the elliptical and *b* is the minor axis of the elliptical. The viewing angle determines what the galaxy looks like. On top of the E0 to E7 classification, there is a giant elliptical galaxy class called *cD*. This class was added later than the others.

The luminosity of normal and giant elliptical galaxies is usually a few multiples of the luminosity of the Milky Way. These elliptical galaxies are generally in the size range of tens of kiloparsecs. The cD class of elliptical galaxies can be in the size range of hundreds of kiloparsecs. The cD systems are also very luminous, as their luminosity can be up to 100 times that of the luminosity of the Milky Way.

2.2.2 Lenticular

The lenticular galaxies are a type of a midway point between spiral and elliptical galaxies and are labelled as S0 galaxies. This type of galaxies contains very little interstellar matter. These galaxies do not have any sort of spiral structure. Lenticular galaxies are more common in the more galaxy rich areas of the Universe. The lenticular galaxies share similarities with both spiral and elliptical galaxies. The lack of a considerable amount of gas and dust lanes is a similarity with the elliptical galaxies. The similarity to the spiral galaxies is in the flat disk that is made up of stars.

2.2.3 Spiral

Spiral galaxies have been named for their spiral pattern. The spiral pattern can be more or less pronounced. The basic build of a spiral galaxy is a central bulge, a stellar disk, and bright spiral arms. Bulge of a galaxy is a spheroidal dense, tightly packed, cluster of stars. The central bulge resembles the structure of an elliptical galaxy. The stellar disk of a spiral galaxy is similar to a lenticular galaxy. Clusters of bright hot O and B stars outline the arms of a spiral galaxy. Compressed dusty gas located inside the arms of the galaxy form the O and B stars. O-type stars are very luminous blue-white stars with a temperature over 30 000 K. B stars are blue or bluish white stars with a temperature between 10 000 K and 30 000 K.

Spiral galaxies are divided into two sequences, barred and ordinary. A central linear bar is present in nearly half of all spiral galaxies. The ordinary spiral galaxies are divided into three classes Sa-Sb-Sc and the barred are also divided into three classes SBa-SBb-SBc. A significant difference between these two spiral galaxy sequences is where the spiral pattern ends. In the normal spiral galaxies, the spiral pattern can extend to the center or end at the inner ring. The inner ring is a star-forming region located outside of the ring that surrounds the center of the galaxy. This ring can also be called the nuclear ring. For barred spiral galaxies the spiral pattern always ends at the central bar. Our galaxy, the Milky Way or the Galaxy, is of type Sb, Sbc or SABbc, an intermediate spiral galaxy class between Sb and Sc, between a barred and a normal spiral galaxy.

2.2.4 Irregular

The irregular galaxy class or sequence is for those galaxies Hubble could not categorize into any of his three categories. The Hubble scheme of categorization, as stated works for the large and bright galaxies, which means that the faint galaxies have been much harder to categorize which is why they have been put into the irregular class. The irregular class can be divided in a few different manners, like to classes Irr I and Irr II for irregular 1 and irregular 2. Of these two galaxy classes, the Irr I galaxies creates a continuation to later types, beyond Sc galaxies. The Irr I galaxies contain a lot of young stars and they are plentiful in gas. Irr II galaxies are also rich in gas, but they contain very little young stars. Irr II galaxies resemble irregular small ellipticals.

Dwarf galaxies can be placed into at least three groups, dwarf irregulars (dIrr), the smallest group of irregular galaxies, as well as spherical or elliptic dwarf galaxies called dSph (dwarf spheroidals) and dE (dwarf ellipsoids). The recently found uFd (ultrafaint dwarfs) galaxies, BCD (blue compact dwarfs), and UCD (ultracompact dwarfs) as other classes of dwarf galaxies.

2.2.5 Hubble's galaxy classification map

A basic map of the Hubble's galaxy classification scheme is presented in Figure 2.1. Three elliptical galaxies are presented, E0, E3, and E6. From Figure 2.1 it is possible to see the circular shape of the E0 galaxy. After the elliptical galaxies, the scheme branches into two sections. In this scheme, the lenticular galaxies are divided into two classes, normal and barred lenticular galaxies. The first galaxy classes on both branches is a lenticular galaxy. S0 is the normal lenticular galaxy class, while the SB0 is the barred lenticular galaxy class. The spiral galaxy classes follow from the lenticular classes. On the normal side is the Sa, Sb, and Sc. On the barred side are the SBa, SBb, and SBc barred spiral galaxy classes. The figure shows on a basic level the form differences between the spiral galaxies. The largest differences can be seen between Sb and SBb. The irregular galaxy class is on the rightmost side. The different irregular galaxy types have not been further specified in Figure 2.1.

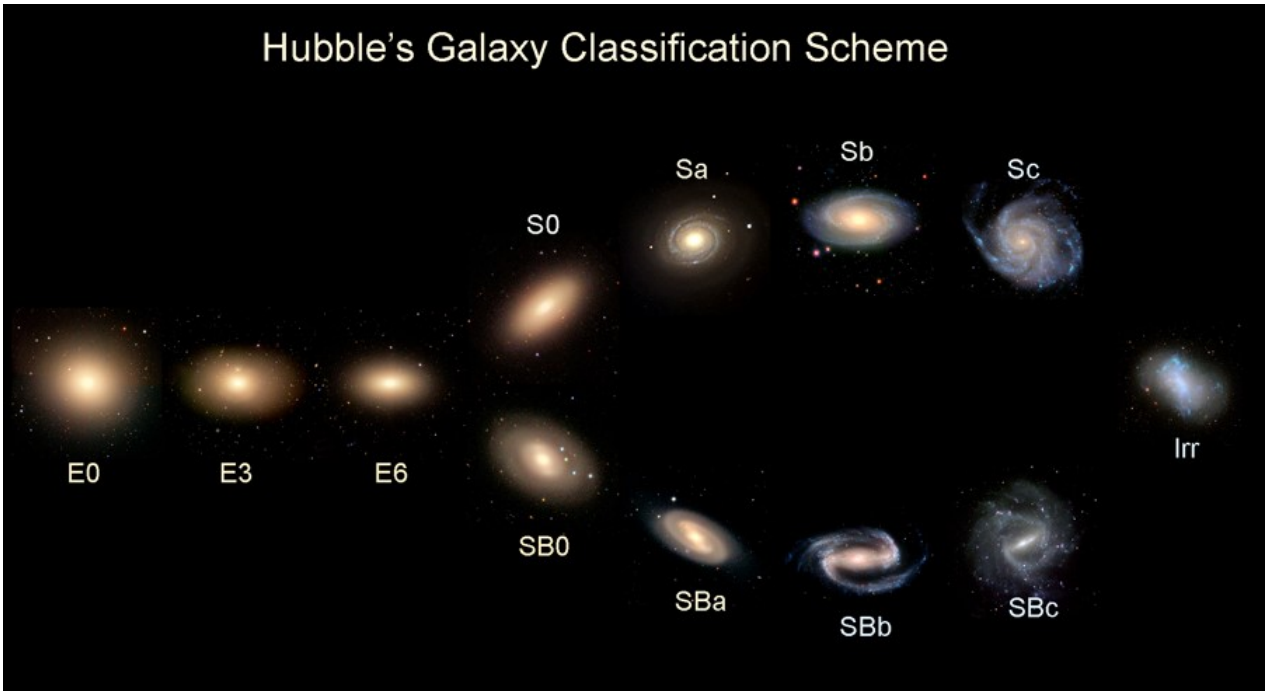

Figure 2.1 Hubble classification scheme of galaxies (Galaxy Zoo)

2.2.6 Other type of classification

There are also other types of classification of galaxies than just classification based on the shape of the galaxy. This type of classification is done when galaxies are plotted into a color-magnitude diagram like stars. Color-magnitude classification has two classes, the red and the blue galaxies. The blue galaxies are generally spiral galaxies with young stars while the red galaxies are usually elliptical galaxies with old stars. In other words, this type of classification is based on an evolutionary condition rather than a morphology condition.

2.3 Galaxy evolution

The galaxies that are monitored today can differ significantly from the galaxies they were millions of years ago. Galaxies and their properties change as they evolve, like for

example the color of the galaxy as well as the morphology of the galaxy. Protogalaxy is a gas cloud that is slowly forming into a galaxy. Scientists believe that active galaxies could be an active evolutionary stage of a nonactive galaxy, like that radio galaxies are an active stage of an elliptical galaxy.

The color of a galaxy can tell about its age as well as other structural properties. In general, it is believed that blue galaxies are younger galaxies (Maddox et al. 1990). Star formation only occurs in a blue galaxy. When a galaxy turns from blue to red, star formation stops. Red and blue galaxies can be found in different type of regions in space. Blue galaxies tend to be in low density regions while red galaxies prefer denser regions in space. The morphology of blue galaxies is usually disk-dominated, while red galaxies frequently have a bulge-dominated morphology (Vulcani et al. 2014).

There are three different ways a galaxy can evolve in: Passive evolution, secular evolution, and through interactions and mergers (Combes 2006). The least eventful of the evolution ways is the passive evolution. If a galaxy undergoes passive evolution, it means that it remains undisturbed from both star formation and collisions with other matter in space. Since there is no new star formation in a galaxy that undergoes passive evolution, the galaxy will slowly star turning red. The color change occurs because the existing blue stars turn into red giants. Passive evolution generally leads to elliptical and S0 galaxies (Vulcani et al. 2014).

Secular evolution is a slow evolution during which a galaxy is undergoing rearrangements in its configuration to become the most tightly bound configuration possible. This occurs when angular momentum or energy is directed outward from either an internal process or an environmental process. The internal process is an induced process driven by for example bar instabilities or bar and spiral structures disbursing both angular momentum and energy that reorganizes the disk structure. Environmental secular evolution is caused by for example small mergers or by galaxy harassment (high-speed galaxy encounters within clusters that occur frequently) (Kormendy 2012).

In an environmentally driven secular evolution, a spiral galaxy that is rich in gas as well as capable of star formation can be transformed into a red (and dead) S0 galaxy or a spheroidal galaxy. This can also happen to irregular galaxies (Kormendy 2013). A secular evolution that is internally driven can cause disk growth for example due to formation and evolution central bars, gas slowly accreting, and spiral instabilities (Sellwood 2013).

Many galaxy properties, e.g. amount of gas, color and morphology, depend on the environment the galaxy is located in. Interactions and merging between galaxies can cause morphological changes (becoming for example elliptical galaxies) and or color changes. As it is possible for the interactions not to cause new star formation, the color does not necessarily change. This type evolution is rather self-explanatory, when a galaxy crosses its path with for example a star cluster or a gas cloud, the properties of the galaxy

change. A spiral galaxy that passes through dense clusters, the galaxy “steals” the interstellar gas of the clusters, thus discarding the clusters ability to form stars (Alonso et al. 2012). Figure 2.2 is a simple visual representation of galaxy evolution processes.

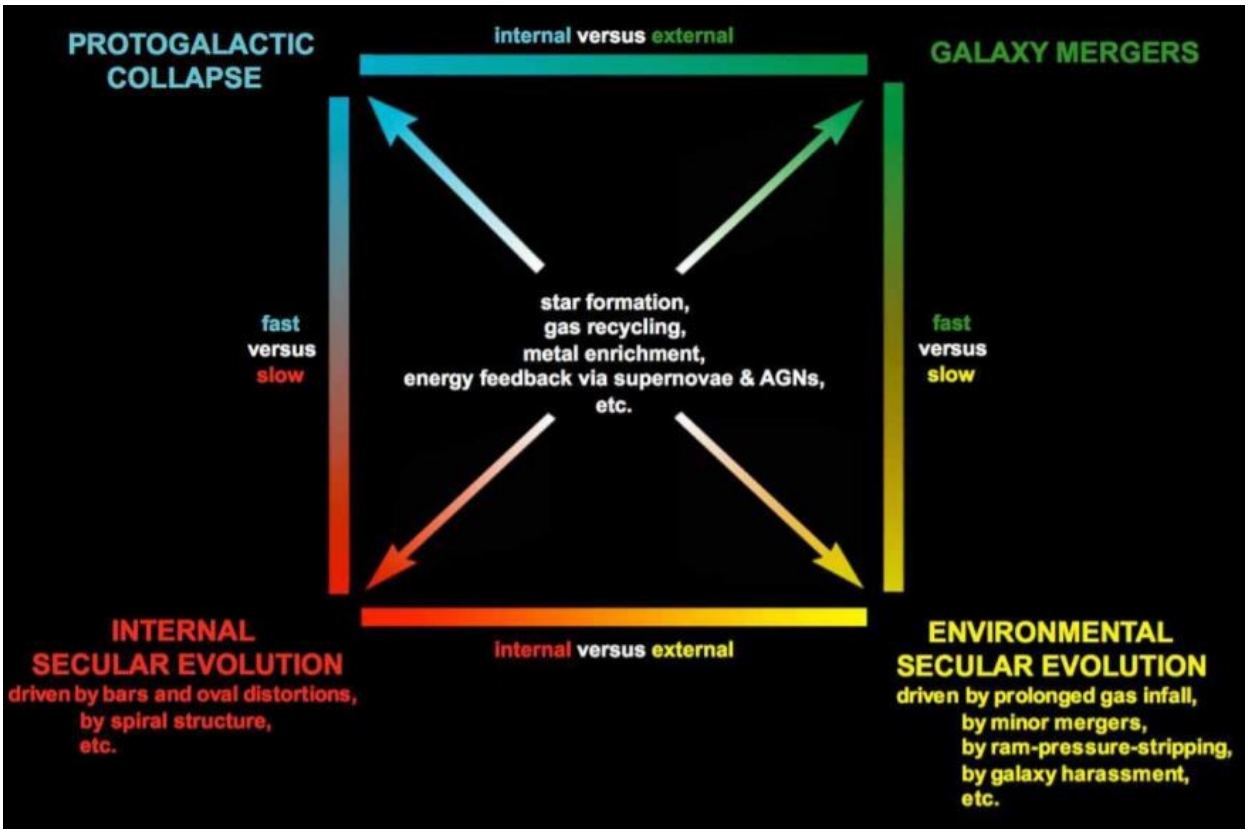

Figure 2.2 Galaxy evolution processes (Kormendy 2013)

3. Active galaxies

Though the name is misleading, in an active galaxy only the core is active, not the entire galaxy. Due to the core being the only active part of the galaxy, many scientists prefer the name active galactic nuclei, or AGN for short. In a regular galaxy, the observed radiation is due to stars and the star heated gas. In an AGN a large portion of the radiation observed is not due to the light of the stars or the star-heated gas, but rather from the area near the central supermassive black hole that accretes mass. According to our current understanding of the Universe, AGNs are the most energetic constant objects we know (Beckmann & Shrader 2013).

Active galaxies emit more energy than the sum of the emission from the stars in a galaxy. The surplus energy in an active galaxy can be found in infrared (IR), radio, X-ray, as well as ultraviolet (UV) regions of the electromagnetic spectrum. Active galaxies are not considered to be a whole new different class of galaxies, but rather a phase in the evolution of galaxies. This is because it is deemed very unlikely for an active galaxy to be able to output so much energy for a long time.

The general structure of an AGN consists of a very bright core, a supermassive black hole in the center of the galaxy, an accretion disk around the core, and a large torus. AGNs can also have relativistic plasma jets. Jets will be further discussed in Chapter 3.1. An illustration of a basic AGN structure can be seen in Figure 3.1 Figure 3.1 has also marked BLR and NLR. Both of these topics will be further discussed in Chapter 3.3.

Compared to the rest of the galaxy, the core in an active galaxy is generally very bright and variable. It is possible for the core to not be more luminous than the rest of the galaxy. This type of active galaxies are often very strong radio sources. There are a few common denominators with these galaxies. One is that active galaxies have a very large energy output. Another common denominator is that the volume of where the energy comes from is very small.

The black hole located at the center of the galaxy has an angular momentum. Gas surrounding the black hole is slowly falling into the black hole. The black hole is surrounded closely by a disk of gas that is called an accretion disk. An accretion disk is formed of plasma, dust, particles, or gas that rapidly rotates around an astrophysical object.

The material in the disk spirals slowly inward due to viscosity, heating up and radiating away its gravitational potential energy. This process ends when it reaches the

last stable orbit around the black hole and finally falls in. The magnetic fields around the black hole are pulled inward with the flow of the hot and ionized gas. At times, this process may cause a jet, as the energy is converted into perpendicular motions along the rotational axis. The jet is then moved outward along the spin axis.

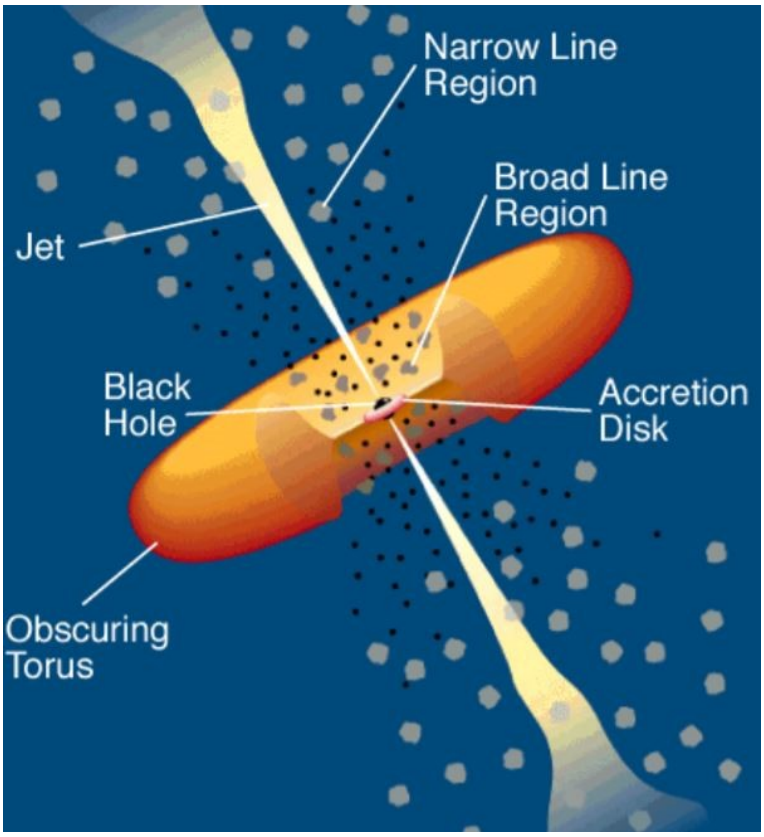

Figure 3.1 A general structure of an AGN (Urry & Padovani 1995).

A torus of an AGN is, as seen in Figure 3.1, very large compared to the accretion disk. The doughnut shaped torus, located 10-100 pc from the black hole, consists of thick molecular clouds, capable of absorbing radiation from the nucleus. The torus is sometimes called dusty absorber (Järvelä 2018).

Active galaxies have many different forms of activity, like brightness of the galaxy. The general belief is that the galaxies with a very bright nucleus are young galaxies, with a large number of stars near the nucleus. These stars evolve and form into supernovae.. The nuclei with the most powerful radiation can outshine the rest of the galaxy. Very luminous active galaxies can also be *quasars*. Quasars can have a luminosity of over

$10^{12} L_{\odot}$. Quasars were much more common when the Universe was younger. Quasars were the most common when the Universe was 20 % - 40 % of its current age.

3.1 Jets of galaxies

Relativistic jets are powerful narrow outbursts of plasma or radiation that originate from deep within the central core of a galaxy. Jets can be superluminal. Jets can be either one-sided or two-sided. Jets show where the energy from the nucleus of the galaxy is being directed. A jet is visualized in Figure 3.1. Two type of relativistic jets exist: galactic jets and stellar jets. Stellar jets have multiple different sources, like T Tauri stars, planetary nebulae, neutron stars, and stellar black holes. Jets can be either one-sided or two-sided. Most jets are found in active galaxies that are strong radio sources, but it is sometimes possible to see a very weak jet in a weak radio source galaxy (Massaglia 2009). Jets can be either optically visible or invisible.

Galactic jets only have one type of source, a supermassive black hole located in the center of a galaxy. Stellar jets are omitted from this thesis, as they are not relevant for it. Not all active galaxies have jets. Jets can be observed in the nucleus of many quasars. Galactic jets can also originate from Seyfert galaxies and radio galaxies. Blazars always have a jet.

Jets consist of electrons as well as other particles. A jet must be neutral, which means the jet must also have protons or positrons to match the electrons. According to the unification model, discussed in Chapter 4, radio brightness determines the brightness and strength of a jet. Accretion disks can cause a relativistic jet when the energy of the disk is converted into perpendicular motion along the rotational axis. In a radio galaxy, jets move the energy from the nucleus of the galaxy to radio lobes. Radio galaxies and radio lobes will be further discussed in Chapter 5.

Astrophysicists have speculated, that narrow outbursts are launched and kept narrow by very strong magnetic fields. These magnetic fields build upward in the surrounding disk of inflowing matter. The outgoing jets may appear to move faster than light. This is, however, an illusion, as the speed at which the jets move, is still slower than the speed of light, though very close to it.

The power of a galactic jet is speculated to originate from the spin and the mass of the supermassive black hole in the center of the active galaxy. The magnetic field near the event horizon is also thought to be a reason for the power of the jet (Ghisellini et al.

2014). Active galaxies can have a nonthermal spectrum. The spectrum is nonthermal due to synchrotron radiation that is produced by fast electrons in a magnetic field. Synchrotron radiation is most commonly observed in the radio spectrum, but sometimes also in the γ-ray band. Due to this, jets can also be called radio jets (La Mura et al. 2014).

3.2 Active galaxy classification

Classification of active galaxies has been relatively non-systematic. This is partially due to active galaxies being a relatively newly discovered phenomenon and many active galaxies have been found recently. There are many different classes of active galaxies, quasars being one the classes of active galaxies. Other classes are for example radio galaxies and Seyfert galaxies. Blazars are also a type of active galaxy.

3.3 Spectral lines of active galaxies

The line emission is one of the most important criteria when classifying active galaxies. In spectroscopy, as by the old definition, the interaction between matter and electromagnetic radiation, whether the radiation is absorbed or emitted, is observed. Nowadays, spectroscopy also includes the interaction between particles like electrons, protons, and ions, as well as observing the collision energy between the particles. Spectroscopy measures the widths, strengths, and positions of spectral lines. The spectral lines can tell different properties of the source or even of the intervening medium.

Spectral lines are either in the NLR (NLR) or the BLR (BLR). BLR and NLR are produced from different locations in the galaxy as seen in Figure 3.1. When observing Figure 3.1, it is possible to see that the BLR is located close to the black hole in the center of the galaxy. The BLR is often located 0.01-0.5 pc from the black hole. The BLR consists of gas clouds that orbit around the black hole with speeds between 1 000 and 10 000 km/s (Järvelä 2018). NLR and BLR are characterized by their widths. The width of the regions is determined separately for each galaxy class (Seyferts, radio galaxies etc.).

Broad-line spectra are characterized by the existence of both high and low excitation lines. The high excitation lines indicate highly ionized material, while the low excitation lines indicate lowly ionized material. The composition of the broad-line clouds

is hard to estimate, but it is possible to determine that due to the relative strength of the lines, there must be a Solar abundance of heavy elements.

The NLR is, as seen in Figure 3.1, located further away from the black hole, accretion disk, and torus, over 100 pc from the black hole (Järvelä 2018). The NLR also consists of gas but unlike in the BLR region, in the NLR, the gas moves with a low velocity. The gas in the NLR region is also low density compared to the gas in the BLR region. This low-density gas causes the forbidden narrow emission lines (Järvelä 2018). Forbidden lines are further discussed in Chapter 3.4.

Gustav Kirchhoff and Robert Bunsen deduced in the mid-19th century, that absorption and emission lines are due to the existence of certain elements that each possess their own unique pattern of lines. Based on their discoveries, Kirchhoff, the man who is well known for his circuitry laws, created his three laws of spectroscopy:

(1) When high density gases and solid bodies are heated, they cause a continuous emission spectrum.

(2) Heating of a low-density gas produces a bright line emission. The chemical composition of the gas, as well as the temperature to which the gas is heated determine the presence and strength of the emission lines.

(3) Gas that is placed in between the observer and a hotter source of continuum radiation, dark absorption lines created by the gas cross the continuum. The strength and presence of the dark absorption lines is determined by the temperature of the intervening gas as well as the chemical composition of the gas.

The work of Kirchhoff and Bunsen was instrumental for the future of modern science. The new atomic theory created by Nils Bohr in 1913, can be argued to be only possible due to the discoveries of Kirchhoff and Bunsen.

The spectral lines of active galaxies are very different than the spectral lines of normal galaxies. For example, the spectral lines of the very first observed active galaxies had a prominent and broad emission lines, caused by highly ionized species of atoms. The spectral lines are useful for many other reasons too than just for identifying active galaxies. When studying spectral lines, it is possible to determine different physical processes that occur in the emission region.

In principle, it is possible to determine the chemical composition of the radiating matter as well as the ionization or excitation of the material. The ionization or excitation of the material relates to the temperature of the matter. The spectrum of a galaxy can

also be used in principle for obtaining information about the velocity field of the atoms, which relates to the temperature of the atoms. The motion along the line-of-sight can also be deduced from a spectrum. Large internal velocities can be due to explosions in the nucleus or due to rotational velocity near a black hole.

3.4 Forbidden lines

In astronomical spectroscopy, forbidden lines are bright emission lines that occur due to certain electron transitions that are normally not permitted by the selection rules of quantum mechanics. Even though the transitions are normally not permitted, there still exists a possibility in which they occur. As it is still possible for the spectral lines to occur, calling them forbidden spectral lines can cause confusion. Shortly put, forbidden lines are a result of a de-exciting atom moving to a lower energy state in a certain environment. An example of a certain environment is a rarefied vacuum of space. Based on a geometrical model, scientists believe that the permitted lines originate from a more dense, inner zone with rapidly moving gas that is adjacent to the central source. The forbidden lines in this model would thus originate from an outer zone of lower density and lower velocity (Dibaj 1981).

The forbidden lines are not always visible, as they can only be seen when the density of atoms is less than the critical density. The density of the gas is important for other reasons too than just for the visibility of the forbidden lines, the density is also important for the emissivity of the gas. For permitted lines, the emissivity of gas per unit mass of gas, increases linearly with the density of the gas. For forbidden lines the emissivity is constant above a critical density. This is because the collisional excitations directly match the collisional de-excitations above the critical density. From this, it is possible to conclude that the ratio of permitted line flux to forbidden line flux increases as the density increases. The higher the density, the harder it is to see the forbidden lines. Forbidden lines are stronger in high-density gases than in low-density gases. However, due to the fact that permitted lines have a bigger increase in flux is much stronger in high-density gases, it is very hard to detect forbidden lines in the high-density gas.

4 Unified model of active galaxies

The unified model of active galaxies is a widely accepted schematic model used to unify the diverse forms of activity in galaxies. In the unified model, the object is to model properties of each individual object, so that it depends on only a few parameters. In this model, majority of galaxies have a compact central core. This core or nucleus is a supermassive black hole, with a mass of $10^6 M_\odot - 10^{10} M_\odot$. In this model, the galaxies also have an accretion disk and a torus.

The total luminosity of a galaxy is the first characteristic parameter of the unified model. This is due to the fact that when for example comparing Seyfert-1 galaxies and radio-quiet quasars, the only essential difference, besides the shape of the host galaxy, is that the radio-quiet quasars are more luminous. Another characteristic parameter that is commonly used in the unified model is radio brightness. In this case radio brightness can be related to the strength of a jet. By observing the radio brightness, it is possible to connect radio galaxies and radio quasars.

The viewing angle of the nuclear disk is another very important parameter. Different viewing angles can change the shape and what we see of a galaxy radically. If the disk is for example viewed edge-on, it becomes impossible to see the core of the galaxy. Issues also arise when viewing a galaxy face-on. If we have a galaxy that resembles a double radio source when viewed edge-on, this same galaxy may appear like a radio quasar when viewed face-on. When viewing face-on, it is possible that the viewing angle happens to be right along a jet, and the galaxy will then appear as a blazar. Figure 3.2 shows an illustration where two different viewing angles are compared.

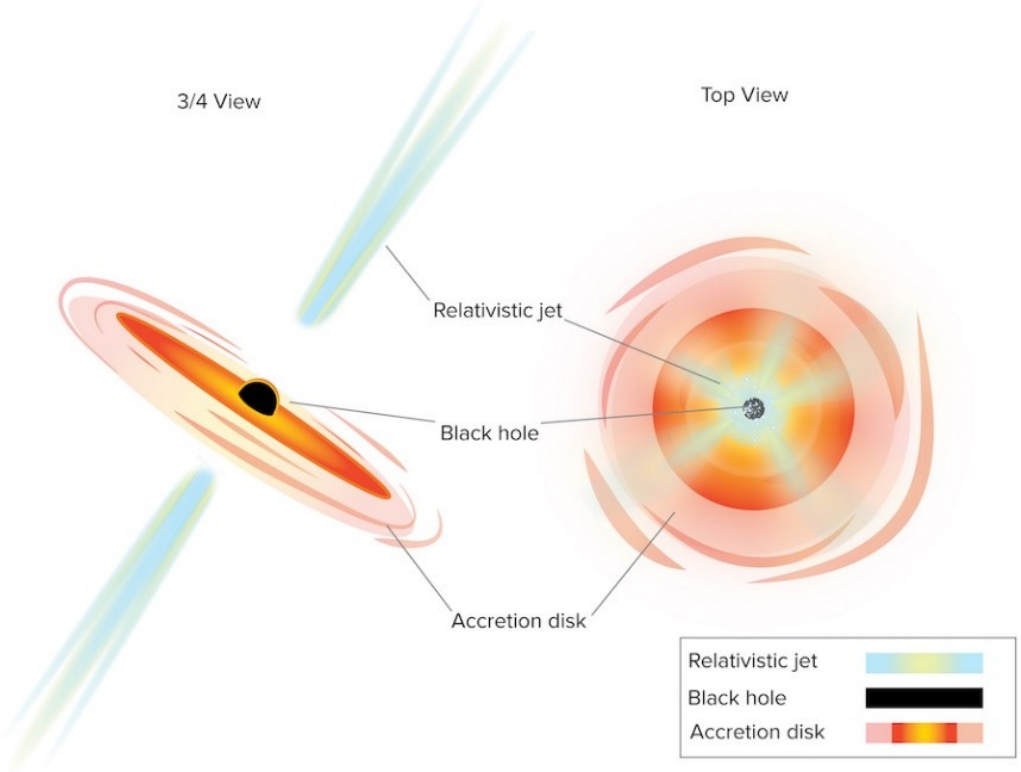

Figure 3.2 A comparison between two viewing angles, ¾ view and top view/face-on view (NRAO/AUI/NSF).

Figure 3.2 compares the viewing angles ¾ view to top view or face-on view. The ¾ view used in this comparison is close to an edge-on view, as the center of the galaxy is barely visible. In the comparison it is possible to see how much the viewing angle changes the visible shape of the galaxy, as well as the location of the black hole, relativistic jet, and accretion disk. In this example, the face-on view appears to be a blazar.

Sources that are viewed face-on are called type-1 quasars or type-1 AGN, while sources with an edge-on view are called type-2 quasars or type-2 AGN. The unified model predicts that there is a large number of quasars with a dusty torus obscuring the nucleus, like in Seyfert-2 galaxies. This type of a quasar is called type-2 quasar. Type-2 quasars only have narrow emission lines (Lyu & Rieke 2017). These types of sources are not included in surveys that are performed in UV, soft X-ray, or optical wavelengths. The surveys that these sources are included in are performed in for example hard X-ray or far infrared wavelengths, as there is far less obscuration in hard X-ray and in far infrared the

absorbed energy is re-radiated. Type-1 quasars are for example Seyfert-1 galaxies and have both broad and narrow emission lines (Lyu & Rieke 2017).

NASA has performed multiple surveys at hard X-ray and far infrared regions, with two different space satellites. NASA used Chanda X-ray satellite and the Spitzer Space Telescope for these surveys. The final result of both surveys was that at least 75 % of all of the supermassive black holes are heavily obscured.

5 Radio galaxies

Radio astronomy was a very popular branch of science in the 1950s. During this time, scientists found numerous strong radio sources. Today many of those sources are known to be radio galaxies. Radio interferometers reached the same resolution as an optical telescope in the 1950s. Due to this leap in science, it was possible to start observing both the size and the form of objects in the sky, which we nowadays know to be radio galaxies.

Radio galaxies are, as the name hints, galaxies that have strong radio emission that originates from relativistic jets. The emission of these powerful radio sources is non-thermal synchrotron radiation. Most of the known radio galaxies are elliptical galaxies. Radio galaxies are divided into two main categories, Fanaroff and Riley type I and type II galaxies, also known as FRI and FRII galaxies. Radio galaxies can have a so-called double structure. A double structure in this case means that there are two large, radio emitting regions that are located on the opposite sides of the galaxy. The double structure is a result of jets from the nucleus of the galaxy. The radio emitting regions are also known as radio lobes.

Two of the main differences between the FRI and FRII galaxies lie in the radio luminosity and brightness of the ends of the radio lobes. The radio luminosity of FRI galaxies is lower than the radio luminosity of FRII galaxies. The difference between the ends of the radio lobes is that for the FRI galaxies, the ends fade out while for the FRII galaxies, the ends as well as the edges of the lobes remain bright.

In a radio galaxy, jets direct the energy from the nucleus of the galaxy outward to the radio lobes, thus “feeding” the lobes. FRI galaxies can have either one-sided or two-sided jets. In most cases, a parsec scaled FRI galaxy has a one-sided jet, while kiloparsec scaled FRI galaxies most commonly have two-sided jets. When observing a FRII galaxy, jets tend to be much harder to detect. Figure 5.1 is a visual representation of a possible structure of a radio galaxy.

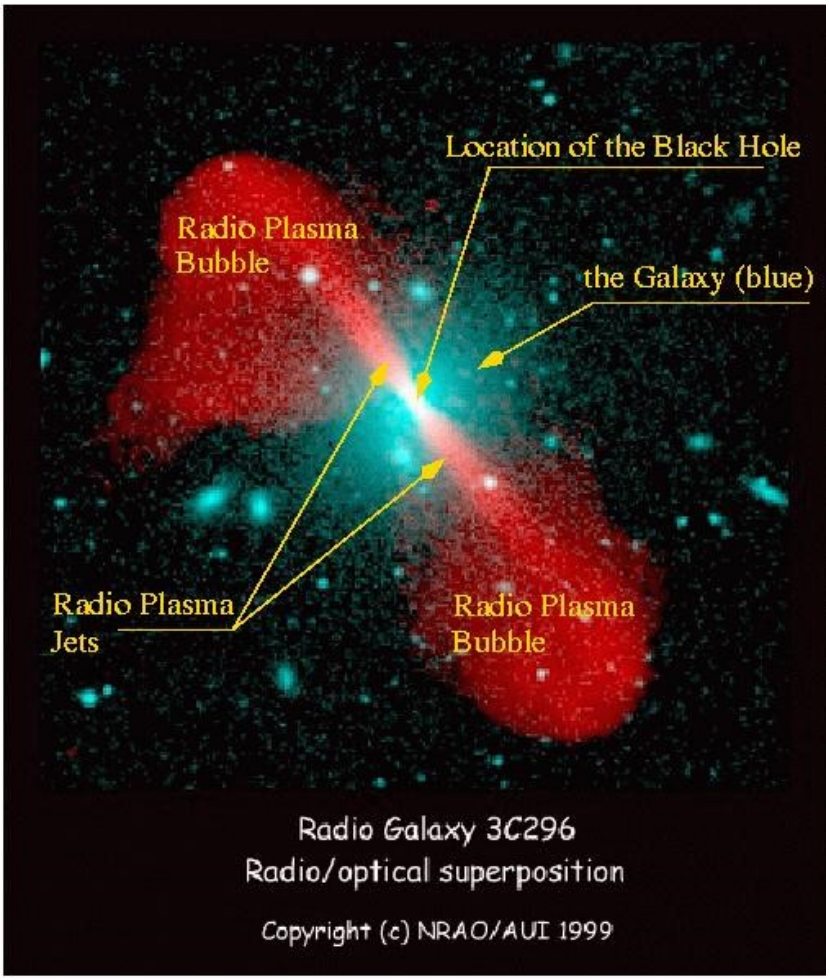

Figure 5.1 A visual representation of a possible radio galaxy structure (NRAO/AUI 1999).

The luminosity of radio galaxies generally varies between $10^7 L_\odot - 10^{10} L_\odot$. This means that the luminosity of a radio galaxy can correspond to the luminosity of a normal galaxy.

The lobes in a radio galaxy are the most luminous at low radio frequencies. These lobes are also optically thin. Within the lobes, there are luminous spots, sometimes called hot spots. Hot spots can emit polarized visible light (Lähteenmäki & Valtaoja 1997). The hot spots are usually located in a symmetrical manner within the lobes, with respect to the nucleus.

The distance between the radio lobes can be as large as 6 Mpc or as small as a few kpc. A stronger radio source results in larger distance between the lobes, while weaker radio sources have a smaller distance between the lobes. This is a result of jets and the jets “feeding” the radio lobes. For the region between the lobes to be filled with emitting material, the nucleus of the galaxy must have been active for a minimum of 10-50 million years. Electrons in the lobes are not a result of ejections from the nucleus, as they would lose all of their energy in the transit from the center of the galaxy to the outer edges of the galaxy. This means that electrons are repeatedly accelerated within the lobes to upkeep their energy.

Radio galaxies with very large radio lobes are, in most cases, the brightest in a galaxy cluster. These galaxies are giant elliptical and cD galaxies. Many of the powerful radio galaxies are blue in color and show recent star formation by the core of the galaxy. The less powerful radio galaxies generally appear very faint in visible light. The emission-line spectrum of a powerful radio galaxy can look similar to that of a Seyfert galaxy.

6 Quasars

"Twinkle, twinkle quasi-star
Biggest puzzle from afar
How unlike the other ones
Brighter than a billion suns
Twinkle, twinkle, quasi-star
How I wonder what you are."
George Gamow

Quasars are a subclass of AGNs and the most luminous long-term objects that we currently know of. Quasars are also the furthest objects we can see. The distance can be determined from the redshift. Supernovae and gamma-ray bursts can be brighter than a quasar, but they are very short-term. Quasars can easily outshine their host galaxy. Quasars can appear like point sources or quasi-stellar. The proper name of a quasar is quasi-stellar radio source. However, scientists prefer a different term, they prefer the abbreviation QSO (quasi-stellar object). The reason why scientists prefer QSO over quasi-stellar radio source is because not all quasars emit radio radiation.

The luminosity of quasars varies greatly and at times very rapidly. Quasar luminosity is mostly between $10^{38}\,\mathrm{W} - 10^{41}\,\mathrm{W}$ or $10^{7}\mathrm{L}_{\odot} - 10^{15}\mathrm{L}_{\odot}$. The luminosity of a quasar can vary very rapidly, within a few days or even in less time than that. The emitting region of a quasar is no larger than a few light-days, or roughly 100 AU.

The first quasar was found in 1963 by Maarten Schmidt. Schmidt was studying the optical emission lines of the known radio source 3C273. Schmidt noted that the hydrogen Balmer lines redshifted by 16 %. A redshift of 16 % is a very large redshift. A large redshift is a characteristic feature of a quasar. The light in quasars is generally polarized only by a few percent. This would indicate, the light in quasars could partially be synchrotron radiation from a relativistic jet. However, not all quasars have jets.

The first quasars were observed through radio observations, but only roughly 10 % of all quasars are strong radio sources. Those quasars with strong radio sources are called radio-loud, the quasars that do not have strong radio sources are called radio-quiet. The radio-loud/radio-quiet can be applied also to other galaxies like Seyfert.

7 Blazars

A blazar is an AGN with a relativistic jet that is directed either totally or very nearly towards the Earth. Blazars are often found in elliptical galaxies. Blazars are objects with very violent and rapid variations in both brightness and polarization. BL Lac sources can also have invisible or very faint emission lines. If the jet is nearly relativistic, so nearly perfectly aligned towards the Earth, the transverse velocity is superluminal. Blazars are related to quasars. Blazars have two main categories: BL Lacertae objects, or BL Lac objects, and flat-spectrum radio quasar (FSRQ). FSRQ was before known as optically violent variable (OVV).

Originally the BL Lac objects and FSRQ quasars were thought to be subclasses of quasars, but in 1978, in Pittsburg, Pennsylvania, scientists came to the end result that these subclasses should create their own class, nowadays known as blazars. The grouping was done because OVV quasars and BL Lac objects had many observational properties in common with each other. Roughly every thousandth quasar is a blazar.

8 Seyfert galaxies

The Seyfert galaxy type is usually divided into two groups, type 1 and type 2 Seyfert galaxy. It is possible to perform a further distinction with fractions, for example type 1.5, type 1.2, and so on. The closer the number is to 1, the more the galaxy resembles a Seyfert-1 galaxy, the closer the number is to 2, the more the galaxy resembles a Seyfert-2 galaxy. Scientists argue that the only difference between type 1 and type 2 Seyfert galaxies is the viewing angle, as discussed in Chapter 4. Type 1 Seyfert galaxies can be also be divided into two categories based on the emission lines. These two categories are narrow-line and broad-line (Osterbrock & Pogge 1985). Narrow-line Seyfert-1 galaxies are discussed in detail in Chapter 9.

One common denominator for all different Seyfert galaxies is that they are in at least majority of cases spiral galaxies. The possible exceptions to all Seyfert galaxies being spiral can be found in both Seyfert-1 and Seyfert-2 galaxies. When Carl Seyfert, the discover of the class, originally published his list of 12 galaxies, 11 out of these galaxies were spiral but one of them, NGC 1275 was an elliptical galaxy, a gigantic cD galaxy. The vast majority of Seyfert galaxies are either spiral or S0 galaxies. About 10 % of all Sa and Sb galaxies have a Seyfert nuclei. A Seyfert nuclei or core is a point-like central nucleus that is very bright and was before believed to always have weak radio emission (Mundell et al. 2009). Weak radio emission is emission that originates from the star formation process rather than from relativistic jets. An estimated 1-2 % of cores of all the known bright galaxies are a Seyfert core. The luminosity of the nucleus of a Seyfert galaxy is between $10^{36}\ \mathrm{W} - 10^{39}\ \mathrm{W}$, so $10^{10}\ \mathrm{L}_{\odot} - 10^{13}\ \mathrm{L}_{\odot}$.

Seyfert galaxies were for a long time thought to be weak radio sources, but according to new research this is not always the case. Scientists have found Seyfert galaxies that show relatively strong radio emission generated by relativistic jets (Komossa 2018; Lähteenmäki et al. 2017; Lähteenmäki et al. 2018). Other important characteristics include very high luminosity, a point-like central nucleus, as well as a spectrum that shows broad emission lines. The emission lines of Seyfert galaxy spectrum are generally believed to be caused by very fast-moving gas clouds that move close to the nucleus. A common trait of Seyfert galaxies is the outward directed flow of matter and energy.

Seyfert galaxies and quasars share similarities, like for example they both shine bright in infrared, visible, X-ray, and ultraviolet wavelengths. Seyfert galaxies are at times thought to be nearby, low-luminosity versions of quasars. The mass of the core in a Seyfert galaxy is in the range of $10^{6}\mathrm{M}_{\odot} - 10^{9}\mathrm{M}_{\odot}$ (Onken & Peterson 2002). The gas clouds around the core that are responsible for the broad emission lines counts for only

a few tens of Solar masses (Dibaj 1981). When comparing a radio-quiet quasar to a Seyfert-1 galaxy, the only essential difference lies in the total luminosity of the two galaxy types. The radio-quiet quasars are more luminous than Seyfert-1 galaxies. The widths of forbidden lines for quasars and Seyfert galaxies correspond to velocities less than $1000\,\frac{\mathrm{km}}{\mathrm{s}}$.

The BLR of a Seyfert galaxy has gas that forms dense clouds with $n_H \gtrsim 10^{10}\,\mathrm{atoms\,cm^{-3}}$. The continuum radiation from the nuclei of the vast majority of Seyfert galaxies is $\lambda < 912\,\text{Å}$. Photons with these wavelengths are very close to the Lyman limit. This means that if the photons would be to travel across the broad-line emitting galaxy, they would be absorbed. Due to this, the clouds must cover only a very small part of the central source. For the NLR, the width of the region is noted to be only a few kiloparsecs across.

The luminosity of a Seyfert galaxy core is so strong that it exerts a considerable amount of pressure onto the gas that surrounds the core. If the pressure grows too high, the interstellar matter cannot fall into the core anymore, causing the galaxy to slowly become non-active.

When observing Seyfert galaxies, it is once again important to take into account the viewing angle as a galaxy can appear like either type 1 Seyfert or type 2 Seyfert based on the viewing angle. An example that visualizes the importance of the viewing angle is presented in Figure 4.1.

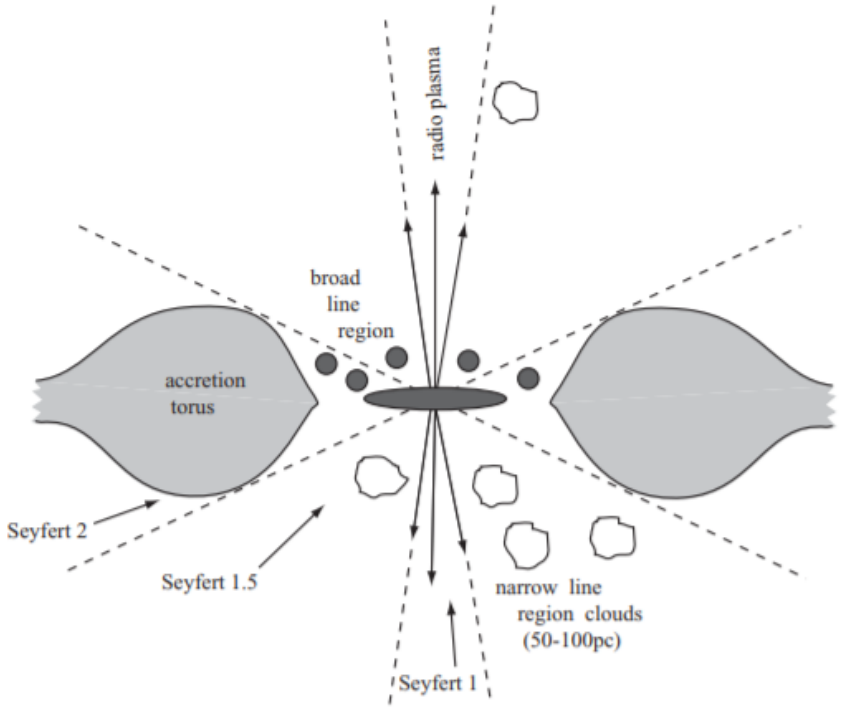

Figure 4.1 A simple visualization of the importance of the viewing angle when observing an active galactic nucleus. The arrowed lines emerging from the center are twin jets. On the sides of the nucleus are two accretion opaque tori. (Sparke & Gallagher 2007)

Figure 4.1 is a very simple model of an active galactic nucleus. When observed face-on, along a jet like in Figure 4.1, the nucleus will appear as a Seyfert 1. If you would, however, observe the nucleus edge-on, along the torus, like in Figure 4.1 the nucleus will appear as Seyfert 2. If one would be to observe the nucleus from a middle point between the two viewing angles, the nucleus will appear as Seyfert 1.5.

8.1 Type 1 Seyfert

The major difference between the two main classes of Seyfert galaxies, type 1 and type 2 Seyfert or Seyfert-1 and Seyfert-2, is the spectrum as well as the viewing angle. For example, the forbidden lines, as well as the core of the permitted lines are narrow for type 1 Seyfert galaxies. The spectrum and viewing angle differences are, however, not the only differences between the classes. Other differences between the two types can

be for example morphological or luminosity differences. This chapter will discuss the basic properties of type 1 Seyfert galaxies.

Type 1 Seyfert galaxies are, are assumed to be, spiral galaxies. As mentioned above, true Seyfert galaxies are often weak radio sources. Seyfert galaxies can, however, be strong sources in other types of radiation. Type 1 Seyfert galaxies frequently emit strong X-ray radiation. Ultraviolet emission is also a common observation when studying type 1 Seyfert galaxies (Malkan & Sargent 1982). Extreme-ultraviolet emission has occasionally been noticed (Vennes et al. 1995). Seyfert-1 nuclei are in most cases more luminous than Seyfert 2 nuclei when observed in wavelengths from infrared to soft X-rays.

The morphology of Seyfert galaxies can heavily depend on the viewing angle, as shown in Figure 4.1. The general rule of thumb is that type 1 Seyfert galaxies are viewed face-on, while type 2 Seyfert galaxies are viewed edge-on. This is strongly supported by the unification model discussed in Chapter 4. Scientists have, however, shown that the morphological differences between the two Seyfert galaxy types are not all simply due to the viewing angle. If that was the case, then the host galaxy and the different types of Seyfert galaxies should remain independent of each other (Chen & Hwang 2017).

Simkin and Mehlberg describe morphological differences between type 1 and type 2 Seyfert galaxies in their article from 1989. One of the differences Simkin and Mehlberg found in their sample, was that when comparing the structure of loops, arcs, and spiral structures in the structure of the non-axisymmetric features of type 1 and type 2 Seyfert galaxies. In type 1 Seyfert galaxies, these arcs, loops, and spiral structures are less diffuse and narrower than in type 2 Seyfert galaxies. This suggests that the new star forming region is more localized for type 1 Seyfert galaxies. The region of H^{2+} activity is also more localized for type 1 Seyfert galaxies. Local starburst activity differences of this sort can potentially reflect differences in the distribution of underlying gravitational potential (Simkin & Mehlberg 1989).

Type 1 Seyfert galaxies are more commonly found within dense galaxy regions (Chen & Hwang 2017). The host galaxy of Seyfert-1 galaxies tend to be more luminous than the host galaxies of Seyfert-2 galaxies (Maia, Machado & Willmer 2003). According to Kolouridis et al. type 1 Seyfert galaxies tend to be closer to larger galaxies than type 2 Seyfert galaxies. Koulouridis et al. also claims that there is no obvious explanation as to why this seems to occur (Koulouridis et al. 2003).

The spectrum of the core of Seyfert-1 galaxies can look similar to that of the spectrum of a quasar. As mentioned earlier, scientists view Seyfert galaxies to be a weaker, less luminous form of quasars. Dibaj states that there is also a possible relation between the mass of the two galaxy types (Dibaj 1981). Both Seyfert-1 nuclei and quasars

have broad emission lines where a wide range of ions can be seen. Ultraviolet photons can excite for example the Balmer lines of hydrogen as well as the lines of singly ionized species, like Mg^{2+}. The ultraviolet photons are also present in the H^{2+} region around hot stars. Others require more energy to be excited.

8.2 Type 2 Seyfert

Seyfert-2 galaxies differ from Seyfert-1 galaxies in a few ways. For example, type 1 Seyfert galaxies have both broad and narrow lines, while type 2 Seyfert galaxies usually only have narrow lines. The width of the spectral lines of Seyfert-2 galaxies are $\lesssim 1000\,\frac{\mathrm{km}}{\mathrm{s}}$. The commonly missing broad line section of Seyfert-2 galaxies can be due to observational selection effect. There have also been exceptions to the missing broad line section, as Seyfert-2 galaxies can also possess a BLR, like in NGC1068. However, this BLR is only visible in polarized light and thus through reflection. Other Seyfert-2 galaxies have also been observed in polarized light, like for example NGC4258.

When observing the spectrum of a type 2 Seyfert galaxy, the forbidden lines and the permitted lines are of an equal, relatively narrow width. Due to the locations where the permitted and forbidden lines originate from, as described in Chapter 3.4, it is possible that Seyfert-2 galaxies would not have the dense inner zone or the fast-moving gas (Dibaj 1981). It is, however, possible that the denser gas in Seyfert-2 galaxies is just obscured.

Another difference can be found in widths of the permitted and forbidden lines. As mentioned earlier, type 1 Seyfert galaxies have significantly broader permitted lines than forbidden lines. The width of the forbidden lines and permitted lines are nearly equal for type 2 Seyfert galaxies. It is possible, that this could mean that the origin region of the forbidden and permitted lines is the same for Seyfert-2 galaxies. Even though the majority of the emission lines of a Seyfert-2 galaxy are of a very similar width, strong lines, like $\mathrm{H}\alpha$, can possess very faint broad wings.

The luminosity of a Seyfert-2 galaxy nucleus is in most cases, lower than the luminosity of a nuclei of a type 1 Seyfert galaxies when observing in spectral regions ranging from infrared to soft X-rays. The overall luminosity of a Seyfert-2 galaxy decreases as the line widths decrease. Meaning, the narrower the lines, the less luminous the galaxy.

When observing the emission of a Seyfert-2 galaxy, a few differences can be spotted when comparing to the emission of a type 1 Seyfert galaxy. These differences include for example the strength of the energetic X-rays. The proportion of highly energetic hard X-rays is larger for Seyfert-2 galaxies compared to Seyfert-1 galaxies. Type 2 Seyfert galaxies also tend to have stronger radio emission than type 1 Seyfert galaxies. The stronger radio emission is because of star formation that occurs in Seyfert-2 galaxies (Lemke, Stickel & Wilke 2000). This radio emission should not be confused with the radio emission from relativistic jets. The radio emission from star formation is weak when observing at high frequencies, while the radio emission from relativistic jets is strong for the same frequencies (Saikia 2018). Unlike Seyfert-1 galaxies, type 2 Seyfert galaxies can be very strong infrared sources.

The neighboring galaxies of type 2 Seyfert galaxies are according to Koulouridis et al. systematically more ionized than the neighboring galaxies of type 1 Seyfert galaxies. Another conclusion drawn by Koulouridis et al. was that type 2 Seyfert galaxies are significantly likelier to have a near neighbor than type 1 Seyfert galaxies (Koulouridis et al. 2003). Koulouridis et al. noticed when observing most of the non-star forming nucleus neighbor galaxies of Seyfert-1 and Seyfert-2 galaxies of their sample that starburst events had occurred far more recently for type 2 Seyfert galaxies, $< 30\ \mathrm{Myr}$ (mega, 10^6, year) than for type 1 Seyfert galaxies, $> 100\ \mathrm{Myr}$ (Koulouridis et al. 2003).

9 Narrow-line Seyfert-1 galaxies

This chapter discusses in detail Narrow-line Seyfert-1 galaxies, or NLS1 galaxies the most important galaxies for this thesis along with their spectral lines. NLS1 galaxies have been for multiple years now, a topic of a lot of attention in the scientific community. A NLS1 galaxy is a subclass of type 1 Seyfert galaxies. The permitted optical emission lines of NLS1s have a moderately narrow width. The full width at half maximum for the $\mathrm{H}\beta$ line (FWHM $\mathrm{H}\beta$) for NLS1s is $\lesssim 2000\,\frac{\mathrm{km}}{\mathrm{s}}$ (Osterbrock & Pogge 1985).

The most commonly used defining criterion of NLS1s is the relative weakness of the emission of $[\mathrm{O\ III}]_{\lambda 5007}$ as well as the width of the broad component the optical Balmer emission lines. The FWHM is also a common defining criterion of NLS1. The relative weakness of $\frac{[\mathrm{O\ III}]_{\lambda 5007}}{\mathrm{H}_{\beta}} \leq 3$ is another common defining criterion (E. Schmidt et al. 2016). $\mathrm{Fe\ II}$ is also a parameter to pay attention to when observing NLS1 (Mathur 2000).

On top of the above-mentioned properties, narrow-line Seyfert-1 galaxies also have strong $Fe\ II$ line intensities and a soft X-ray continuum. The central massive black hole (MBH) mass (M_{BH}) of a NLS1 is stated by Decarli et al. to be lighter than expected. The expected value comes from known MBH-host galaxy scaling relations. NLS1s have a larger accretion rate on to the MBH than the general accretion rate of a broad-line type 1 Seyfert (Decarli et al. 2008). The soft X-ray slopes of NLS1 are very steep compared to the soft X-ray slopes of broad-line Seyfert-1 galaxies. The X-ray slopes can at times even have very large variability in amplitude for NLS1. NLS1 also possess very steep hard X-ray slopes (Mathur 2000).

When observing NLS1 in either the UV or the optical range, a weak big blue bump (BBB) can be seen. This bump is most likely caused by a shift toward higher energies (Mathur 2000). A BBB is a bump in the spectrum with a rise of over $1\ \mu\mathrm{m}$ and a cutoff energy below $0.6\ \mathrm{keV}$ (Zhou et al. 1997). The BBB of an NLS1 has a very high-energy tail with soft X-ray excess that is both unusually strong and steep Strong IR emission as well as high polarization is also typically present in NLS1s (Mathur 2000).

Many of the above-mentioned properties of the continuum can be explained by comparing the high accretion rate to the Eddington mass limit. The Eddington mass limit is $m = \frac{M}{M_{\mathrm{Edd}}}$ (Mathur 2000). The Eddington mass limit is a theoretical upper boundary of a celestial body's or accretion disk's mass (Gregersen 2011). The Eddington limit can also be thought of in terms of luminosity. In this case, the Eddington luminosity is a

maximum possible luminosity of an accreting celestial body in spherical symmetry (Yoshida 2002).

Eddington ratio is the ratio where the luminosity of the body is compared to the Eddington luminosity. The average Eddington ratio of a narrow-line Seyfert-1 galaxy is much larger than the average Eddington ratio of a broad-line Seyfert-1 galaxy, even by an order of magnitude. The average Eddington ratio of a NLS1 is roughly 1, while the average Eddington ratio of a BLS1 is roughly 0.1 (Decarli et al. 2008).

Scientists speculate that NLS1s are AGNs in an early phase of the evolution of AGNs. This type of theory is for example supported by Mathur (2000). He has four arguments for why NLS1s could be an early phase of an AGN. The arguments are presented below.

Argument 1: The first argument relates to the small black hole mass. Smaller black hole masses are common in small spheroid galaxies. Due to NLS1s accreting very close to the Eddington ratio, the black hole mass should increase rapidly.

Argument 2: The second argument is about super-solar gas phase metallicities. NLS1s have evidence of containing super-solar gas phase metallicities. These metallicities in NLS1s could possibly represent a very early stage in galactic evolution during the rapid star formation in the nucleus.

Argument 3: The third arguments points out how NLS1s have high IR luminosity, something common to young, star-forming galaxies.

Argument 4: The final argument is an analogy with high-redshift quasars. There are those who believe that the high-redshift quasars are at an early evolutionary phase. Based on this, Mathur states it is possible that NLS1s could thus be an early evolutionary phase.

9.1 Spectrum of a NLS1

This subchapter explains properties of the spectrum of a NLS1. According to the optical spectroscopy classification, roughly 15 % of all Seyfert galaxies are NLS1 (Schmidt et al. 2018). NLS1 spectra can be modelled either with a Lorentzian or a Gaussian profile. The use of Lorentzian profile for modelling the spectral lines of a NLS1 is a little bit more uncommon, but Goad, Korista, and Ruff (2012) state that the Gaussian

profile is better for spectrums with a broader FWHM lines. Lorentzian profiles are according to the trio better for spectra with narrower lines. Examples of NLS1 spectra can be seen in Figure 10.1. Figures obtained from Schmidt et al. (2018).

When observing a spectrum and wondering if the observed spectrum is of a NLS1 or not, a few different parameters are observed, as well as the profile of the spectrum. A NLS1 should show a narrow spike (narrow line), while a BLS1 shows a wider spike (broad line). The narrow spike is the Hβ line. If a wide spike is seen in a spectrum, it is with very high odds, not a NLS1. The width of the spike is however not the only thing to observe. As mentioned in the beginning of Chapter 10, the common parameters to observe include Fe II and [O III].

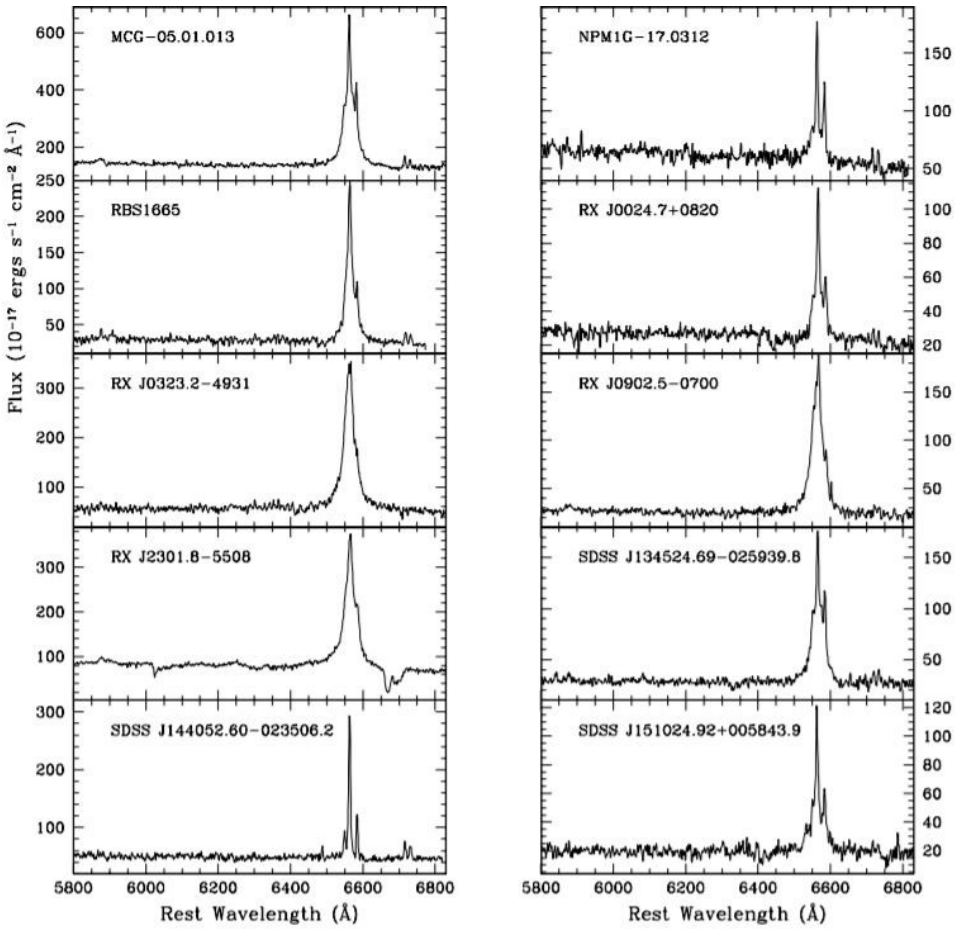

Figure 10.1 Example spectra from different NLS1 sources. The x-axis range is from 5800 Å to 6800 Å, rest frame. The y-axis has flux. Flux is measured in the CGS unit system. The source of each spectrum is seen in the top left of the spectrum (E. Schmidt et al. 2018).

In Figure 10.1 ten different NLS1 spectra are seen. The common denominator for all of the different spectra is that they all have one very clear spike. The $\mathrm{H}\alpha$ line spike is located roughly in the same place for all different galaxies, around 6600 Å that is. The spike for all of these galaxies is narrow, as it should be for a NLS1. A BLS1 would have a much wider spike.

9.1.1 Broad-line Seyfert-1

This is a very brief chapter that summarizes the differences between the subclasses (narrow-line and broad-line) of type 1 Seyfert galaxies. These differences were mostly mentioned in Chapter 9 “Narrow-line Seyfert-1”, though one is mentioned in this chapter. For improved readability, these results are visualized in a table format, in Table 9.1.

As mentioned earlier, in Chapter 4, galaxy classification is based on the point of view we observe the galaxies from. According to this model, NLS1s are viewed at an angle nearly face-on, while BLS1s are viewed from higher angles (Decarli et al. 2008).

Table 9.1 Differences between NLS1 and BLS1

Parameter	**Narrow-line Seyfert-1**	**Broad-line Seyfert-1**
FWHM $H\beta$	$\lesssim 2000\,\frac{\mathrm{km}}{\mathrm{s}}$	$\gtrsim 2000\,\frac{\mathrm{km}}{\mathrm{s}}$
M_{BH}	Smaller than the mass of a BLS1, less than $10^8 \mathrm{M}_\odot$ kg	Larger than the mass of a NLS1, more than $10^8 \mathrm{M}_\odot$ kg
Average Eddington ratio	$\cong 1$	$\cong 0.1$

10 The sample and method of analysis

The data for this thesis has been obtained from *VizieR*[1], a website ran by the University of Strasbourg. This data has the Narrow-line Seyfert 1 galaxies from SDSS-DR12 (Sloan Digital Sky Survey – Data Release 12) as published by Rakshit et al. in April 2017. This data set has a total of 11001 sources. According to Rakshit et al. all of these sources are NLS1. The object of this thesis is to determine which of these 11001 sources actually are NLS1 galaxies. The reason why Rakshit et al. ended up with such a bad sample is because they used a fully automatic analysis process. In spectral analysis, the usage of fully automatic process is usually a bad choice as it often provides spurious results. To achieve the object of this thesis, many different programs and criteria have been used. The programs used for data analysis for this thesis are: Microsoft Excel, Python 2.7, Command Prompt, MATLAB, and Notepad.

Different parameters were obtained from *VizieR* for further analysis. These parameters are: Source ID (SDSS identifier), right ascension, declination, redshift, logarithm of the nuclear monochromatic luminosity at 5100 Å, broad $\mathrm{H}\beta$ component flux, broad $\mathrm{H}\beta$ component FWHM, total (broad and narrow) $[\mathrm{O\ III}]_{\lambda 5007\text{Å}}$ flux, and optical $\mathrm{Fe\ II}$ strength that is relative to the $\mathrm{H}\beta$ broad component.

As mentioned in Chapter 9, NLS1 galaxies can be modelled either with a Gaussian fit or with a Lorentzian fit. Of these two fits, the Gaussian fit is better for a spectrum with broader FWHM $\mathrm{H}\beta$lines and the Lorentzian fit is better for a spectrum with narrower FWHM $\mathrm{H}\beta$ lines. A Python 2.7 code, written by Marco Berton, was used to determine the goodness of fit for both type of profiles for each source. The code can be found in Appendix – Marco Berton's source analysis code. This code would study the spectrum of each source separately and add both a Lorentzian and a Gaussian curve onto the spectral image. The code also saved an image of the spectra and the added fits into a.png file, making it possible to examine the images with basically any image handling software. Figure 10.1 shows an example of an obtained spectrum.

[1] http://vizier.cfa.harvard.edu/viz-bin/VizieR?-source=J/ApJS/229/39

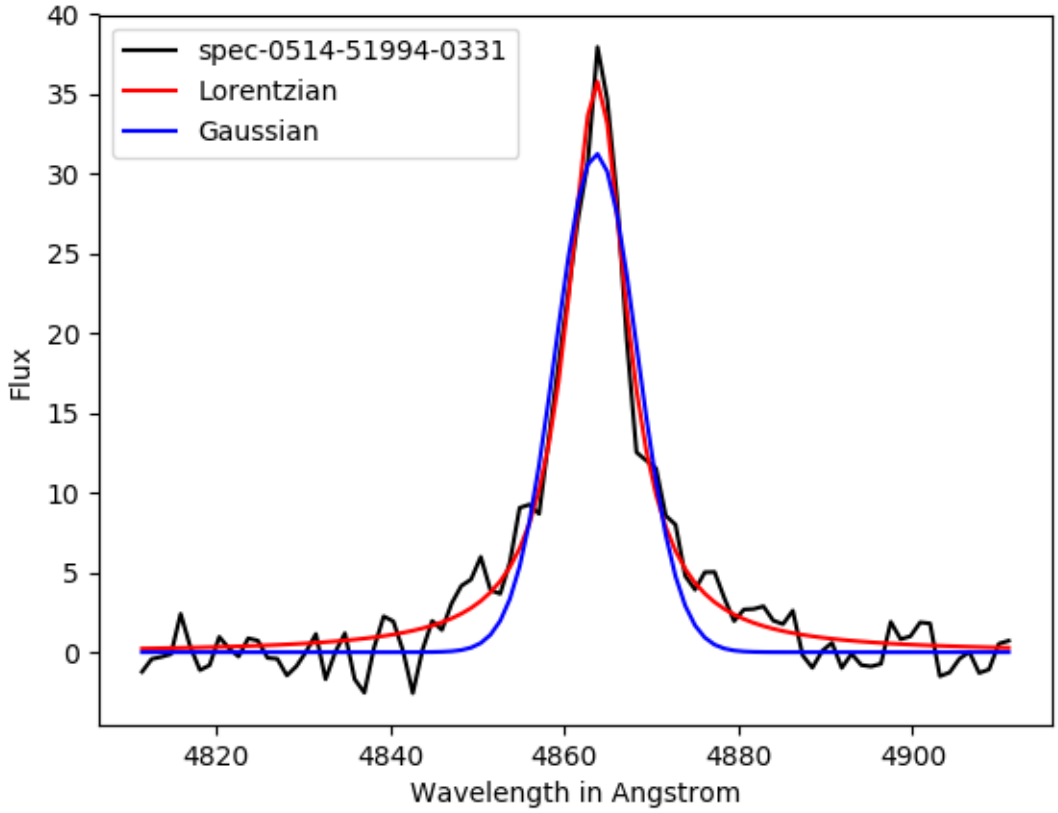

Figure 10.1 An example spectrum obtained with the code by Berton.

Figure 10.1 is an example of a spectrum from a galaxy source. In the figure, Lorentzian profile (red line) can be seen to offer a very good fit and that the Gaussian profile (blue line) fit is not nearly as good. This spectrum does not have that much noise in it, making it easy to read. The code written by Berton also determined other important parameters for each source. The most important parameters obtained with the code are: S/N (Signal-to-Noise) ratio, redshift, RMS continuum, and reduced chi-squared statistic for both Lorentzian and Gaussian profiles.

The reduced chi-squared statistic is a parameter for determining how good of a fit the used profile is. The symbol used for reduced chi-squared statistic is χ^2. If χ^2 is very close or equal to 1, then the fit is good. If $\chi^2 \gg 1$, then the fit is not good. $\chi^2 > 1$ means that the model used has not fully captured the data. When χ^2 is less than 1, it means that model used is over-fitting the data, by for example fitting noise.

All the parameters from the code were saved into a text file. The data from the text file was later transferred to Microsoft Excel. The code provided by Berton also plotted the spectrum for each source.

Microsoft Excel was the software of choice for performing most of the needed mathematics and for organizing of the data. After obtaining all the parameters from the first code provided by Berton, the data needed to be organized into two categories: 1) Most likely NLS1 and 2) Most likely not NLS1. This was done by introducing different discarding criteria. The different criteria used are listed in Table 10.1 as well as the reasoning for why the criteria were chosen.

Table 10.1 Used discarding criteria

Criteria	**Reason why**
S/N ratio below 5	Sources with a lower S/N ratio than 5 have too much noise to be able to determine what they are.
χ^2 value of one fit must be at least below 5 but above 0.5	Sources with a χ^2 value above 5 were deemed to have a very bad fit and sources with χ^2 value below 0.5 over-fitted the data too much and were thus deemed not certainly NLS1.
If the Lorentzian χ^2 value is above 0.5, the Gaussian χ^2 must be below 3	This criterion was used because the Gaussian fit model required a stricter χ^2 upper threshold than the Lorentzian fit model for having similar certainty level to the source being a NLS1.
For a Lorentzian χ^2 value between 3.5 and 5, the S/N ratio must be over 20	This criterion was used to minimize fitting noise for the sources with a high Lorentzian χ^2 value.
For a Lorentzian χ^2 value between 2.5 and 3.5, the S/N ratio must be over 15	Same reason as for the above criterion, but due to the lower χ^2 value, it was possible to lower the S/N ratio threshold.

For a Lorentzian χ^2 value between 2 and 2.5, the S/N ratio must be over 10	Due to a once again lower χ^2 value, it was possible to use a lower S/N ratio threshold.
For a Gaussian χ^2 value between 3 and 5, the S/N ratio must be over 20	This criterion was used to minimize fitting noise for the sources with a high Gaussian χ^2 value.
For a Gaussian χ^2 value between 2 and 3, the S/N ratio must be over 15	The reason for this criterion is as for the above one, but due to a lower χ^2 value, it was possible to use a less strict S/N ratio threshold.
No S/N ratio restriction for Lorentzian and Gaussian χ^2 values between 0.5 and 2	These χ^2 were deemed as very good fits, thus not needing a S/N ratio threshold.

The criteria listed in Table 10.1 were not the original criteria but became the used criteria after the original criteria was determined not strict enough. One of the original criteria was that sources with a S/N ratio below 3, not 5, were discarded. The reason as to why the criteria went up to 5, is because sources with a S/N ratio close to 3, were still deemed to have too much noise.

The used χ^2 thresholds were chosen by inspecting the spectra by eye. The original χ^2 thresholds were also much higher than the final ones. Originally, χ^2 values as high as 50 were going to be used. After observing the spectrum of many sources with χ^2 value above 5, it was noted that the ones above 5, cannot be safely said to be NLS1. When observing the χ^2 values and the spectrum of the sources, it was noted that the Gaussian χ^2 values had to have more strict criteria than the Lorentzian χ^2 values.

Once these criteria were met, the Rakshit et al. measurements were deemed trustworthy. After using all the discarding criteria shown in Table 10.1, 3998 sources remained, 7001 were discarded, and two sources did not run with the code. This means that according to our criteria 63.6 % of the sources claimed by Rakshit et al. to be NLS1, are most likely not NLS1 galaxies. NLS1 galaxies were also lost in the discarding process due to very strict discarding criteria.

After obtaining 3998 saved sources that are most likely NLS1 galaxies, these sources were split into two categories: 1) Gaussian is the better fit and 2) Lorentzian is the better fit. This splitting was done by calculating the absolute value of the expression $1-\chi^2$ for each fit. Then an "If" statement was used in Excel. This "If" statement determined which of the two absolute values was smaller, and thus the better fit. This give 2961 sources with the Lorentzian fit being the better of the two fits. 1037 sources had a better Gaussian fit. For roughly 75 % of the samples, the Lorentzian fit was the better fit. The list of sources can be found from the Metsähovi thesis data database. [2] In the file the Gaussian sources can be found in the section titled 'Appendix A – Gaussian sources'. Lorentzian sources are in section 'Appendix B – Lorentzian sources' of the file. The discarded sources that ran with the code can be found in section 'Appendix C – discarded sources' of the file. The sources that were deemed to have a better Lorentzian fit are very surely NLS1 galaxies, while the Gaussian fitted sources are quite surely NLS1 galaxies. Nothing specific can be stated for the discarded sources.

For better understanding of the sources and possible relations between the two fit types, different parameters were calculated for each source. The calculated parameters are: Black hole mass, bolometric luminosity, Eddington ratio, $[\mathrm{O\ III}]$ luminosity, $\mathrm{H}\beta$ luminosity, and $\frac{\mathrm{H}\beta}{\mathrm{M_{BH}}}$ ratio. The luminosity value for the $\mathrm{H}\beta$ flux and $[\mathrm{O\ III}]$ flux was calculated by using a second code written by Berton. This code was also written in Python 2.7. Due to library problems, Python 2.7 was ran through Command Prompt. This code can be found in Appendix – Marco Berton's luminosity code.

The rest of the parameters were calculated with Microsoft Excel. The equations used for determining the parameters are acquired from the article written by Rakshit et al. in April 2017. The black hole mass was determined by the use of Equation 10.1, or Equation 6 in the article by Rakshit et al.

$$M_{\mathrm{BH}} = \frac{f R_{\mathrm{BLR}} \Delta v^2}{G} \tag{10.1}$$

In this equation, *f* is a scale factor that is determined by Rakshit et al. to be $\frac{3}{4}$. The scale factor is dependent on the geometry and distribution of the BLR. The reason why $f = \frac{3}{4}$, is because the distribution of the BLR clouds is spherical. Δv stands for the FWHM emission line. Gravitational constant is denoted by *G*. $G = 6.67 * 10^{-11}\,\frac{\mathrm{m^3}}{\mathrm{kg{*}s^2}}$ in SI units or $G = 6.67 * 10^{-8}\,\frac{\mathrm{cm^3}}{\mathrm{g{*}s^2}}$ in CGS units. All the math completed in this thesis is done in the CGS

[2] http://www.metsahovi.fi/AGN/thesis-data/Bj%C3%B6rklund_MSc_thesis_Appendices.pdf

system. The last variable of Equation 10.1 is BLR radius (R_{BLR}). The radius is determined by the help of Equation 10.2 or Equation 7 in the article by Rakshit et al.

$$\log\left(\frac{R_{\mathrm{BLR}}}{\mathrm{lt-day}}\right) = K + \alpha * \log(\frac{\lambda L_{\lambda}(5100\ \text{Å})}{10^{44}}\frac{\mathrm{erg}}{\mathrm{s}}) \tag{10.2}$$

K and α are constants given by Rakshit et al. to be $K = 1.527$ and $\alpha = 0.533$. The log expression on the right side of the equation is the division of the monochromatic luminosity at 5100 Å and ten to the 44th power. This value can be calculated in two relatively straight forward ways. As a remainder, the monochromatic luminosity value from *VizieR* is the logarithm value. The longer way of the two ways is to first convert the logarithmic monochromatic luminosity value to the linear value, divide this value by 10^{44} and then take the logarithm of the quotient. The shorter way to determine the right-side logarithm of Equation 10.2, is to utilize basic logarithm rules, by which you simply subtract 44 from the value acquired from *VizieR*.

On the left-side of Equation 10.2, is the BLR radius that is divided by $\mathrm{lt-day}$ which stands for light-day. A light-day is equal to $2.59 * 10^{13}\ \mathrm{m}$ or $2.59 * 10^{15}\ \mathrm{cm}$. Once again, the CGS system is used, meaning the centimeter version of light-day is used.

Equation 10.2 can be transformed to the form shown in Equation 10.3. In Equation 10.3 the radius of the BLR is being solved for, rather than the logarithm expression of the quotient of the radius of the BLR and light-day.

$$R_{\mathrm{BLR}} = 10^{K+\alpha*\log\left(\frac{\lambda L_{\lambda}(5100\ \text{Å})\mathrm{erg}}{10^{44}\quad \mathrm{s}}\right)} * (\mathrm{lt-day}) \tag{10.3}$$

With BLR radius finally solved for, it was possible to determine the black hole mass. To give a better understanding of the black hole mass, the mass was compared to the Solar mass ($\mathrm{M}_{\odot} = 1.989 * 10^{30}\ \mathrm{kg} = 1.989 * 10^{33}\ \mathrm{g}$) by dividing the black hole mass of the source with the Solar mass.

Eddington ratio was solved with in-between steps too. The first needed parameter for determining the Eddington ratio was the bolometric luminosity. The bolometric luminosity was solved by using Equation 10.4, which can also be found in the Rakshit et al. article.

$$L_{\mathrm{bol}} = 9 * \lambda L_{\lambda}(5100\ \text{Å})\frac{\mathrm{erg}}{\mathrm{s}} \tag{10.4}$$

Equation 10.4 is very straightforward as it is a simple multiplication. The only thing that needs to be noted is that the monochromatic luminosity is the linear monochromatic luminosity and not the logarithmic one. The logarithmic bolometric luminosity value is also

an analyzed parameter in this thesis. With the bolometric luminosity, it is possible to solve for Eddington luminosity by using Equation 10.5, also found in the article by Rakshit et al.

$$L_{\mathrm{Edd}} = 1.3 * \frac{10^{38} M_{\mathrm{BH}}}{\mathrm{M}_{\odot}} \frac{\mathrm{erg}}{\mathrm{s}} \tag{10.5}$$

This equation does not introduce any new variables or constants for this thesis, making it very easy to solve. From the Eddington luminosity, it is finally possible to obtain the Eddington ratio, as seen in Equation 10.6.

$$\xi_{\mathrm{Edd}} = \frac{L_{\mathrm{Bol}}}{L_{\mathrm{Edd}}} \tag{10.6}$$

After obtaining each of the different parameters, the mean, median, and standard deviation were calculated with MATLAB. The last thing completed with Microsoft Excel was gathering all of the significant parameters of both profiles in one file for performing the Kolmogorov-Smirnov test with both MATLAB and Python 2.7. The Kolmogorov-Smirnov test or KS- test is a test that tries to determine the similarity or difference between two sets of data. KS- test does not make an assumption based on the distribution of the dataset.

The p-value obtained in the KS- test determines if the *null* hypothesis is rejected or not by observing if the p-value less than the level of significance. In this thesis, a significance level of 1 % is used. This means that the *null* hypothesis is rejected whenever the p-value is below 0.01. The assumption the *null* hypothesis makes is that there is no difference between the two distributions. A rejected *null* hypothesis means that the distributions are different.

The significant parameters for this thesis are: Redshift, R4570, logarithm of the nuclear monochromatic luminosity, Hβ flux, Hβ luminosity, [O III] flux, [O III] luminosity, FWHM, logarithm of the quotient of the division of the black hole mass and the Solar mass, logarithmic bolometric luminosity, and Eddington ratio. A histogram was plotted of each of these 12 significant parameters. The histograms were plotted with MATLAB.

11 Results

This chapter shows the results obtained with methods described in Chapter 10. Most of the results are shown in a histogram form. Each histogram will be briefly discussed as a separate piece for both the Lorentzian and the Gaussian distributions. The general finding for each histogram for each fit will be collected in a table that can be found at the end of the chapter. The findings from this chapter will be discussed in Chapter 12. To make the histograms easier to read, certain histograms have omitted outlier values very far from the rest of the histogram. The omitted values do not change the shape or general distribution type of the histogram. All values have been used for determining the mean, median, and standard deviation. The KS- test also uses all values obtained. The order of the histograms is arbitrary. The first histogram of this chapter is the redshift distribution histogram. All the distribution profile types have been determined by observing by eye, no extra calculations were performed for determining the distribution type.

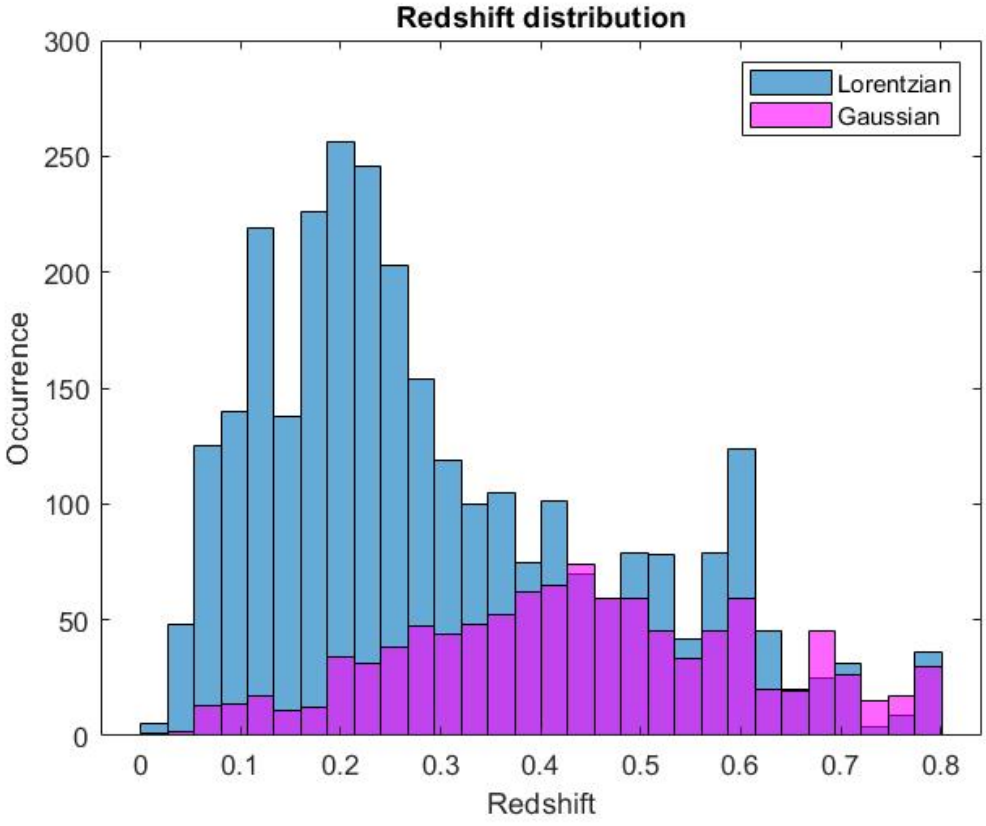

Figure 11.1 Distribution of the redshift for each of the saved sources. Lorentzian sources are marked with blue, while Gaussian sources are marked with magenta. Redshift is marked on the x-axis, while the occurrence is marked on the y-axis.

The redshift distribution, shown in Figure 11.1, for the Lorentzian profiles is a left-sided distribution with no clear distribution profile type. The distribution of the Gaussian profiles is similar to that of a normal distribution, though the very right side of the distribution does not obey the normal distribution. The values with the most occurrences are between 0.19-0.22 for the Lorentzian profiles. As seen in Figure 11.1, there are occurrences of high redshifts for the Lorentzian profiles. At high redshifts, you can lose the wings of the Lorentzian profile.

The Gaussian profiles do not have a very clear most common redshift, but the most common redshifts for the Gaussian profiles are the redshifts between 0.4 and 0.42. The Gaussian profile redshift mean is 0.441, median is 0.4366, and the standard deviation is 0.1738. For the Lorentzian profiles, the mean is 0.3003, the median is 0.2501, and the standard deviation is 0.1783. The p-value obtained from KS- test result for the two datasets is $3.72 * 10^{-102}$. As this value is less than 0.01, so less than the significance level, it means that the *null* hypothesis is rejected, and the distributions are not the same.

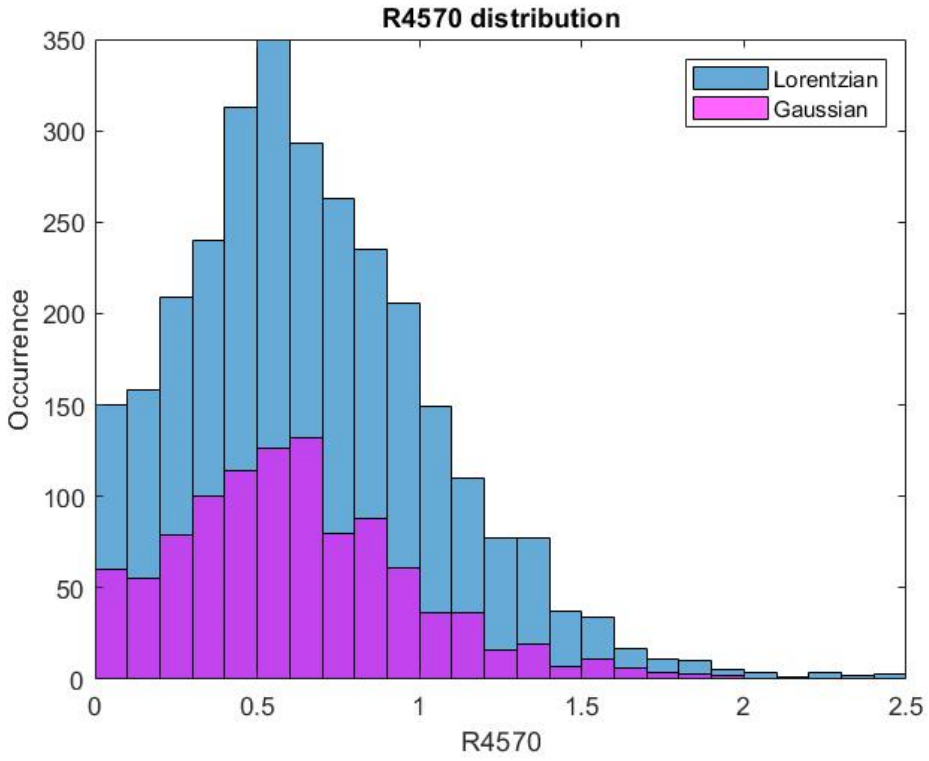

Figure 11.2 Distribution of the R4570 or the optical $Fe\ II$ strength relative to the broad $H\beta$ component. The optical $Fe\ II$ strength is marked on the x-axis while the occurrence is marked on the y-axis. The Gaussian sources are marked with magenta, while the Lorentzian sources are marked with blue.

The distribution of R4570, seen in Figure 11.2 is relatively similar for the two different profile types. R4570 stands for the optical $\mathrm{Fe\ II}$ strength relative to the broad $\mathrm{H\beta}$ component, meaning that this parameter shows the emission strength of $\mathrm{Fe\ II}$ as a function of the broad $\mathrm{H\beta}$ component. Both profile types show a left-sided distribution, showing a significant resemblance to the Rayleigh distribution. The maximum, so the values with the most occurrences, for both distributions lies between 0.5 and 0.7. The mean of the R4570 values for the Gaussian profile is 0.6237, the median is 0.59, and the standard deviation is 0.3699. For the Lorentzian profile, the mean is 0.6704, the average is 0.63, and the standard deviation is 0.4029. The p-value obtained from the KS-test is 0.001035. As the significance level used is 1 %, this means that the *null* hypothesis is rejected, which means that the distributions are different.

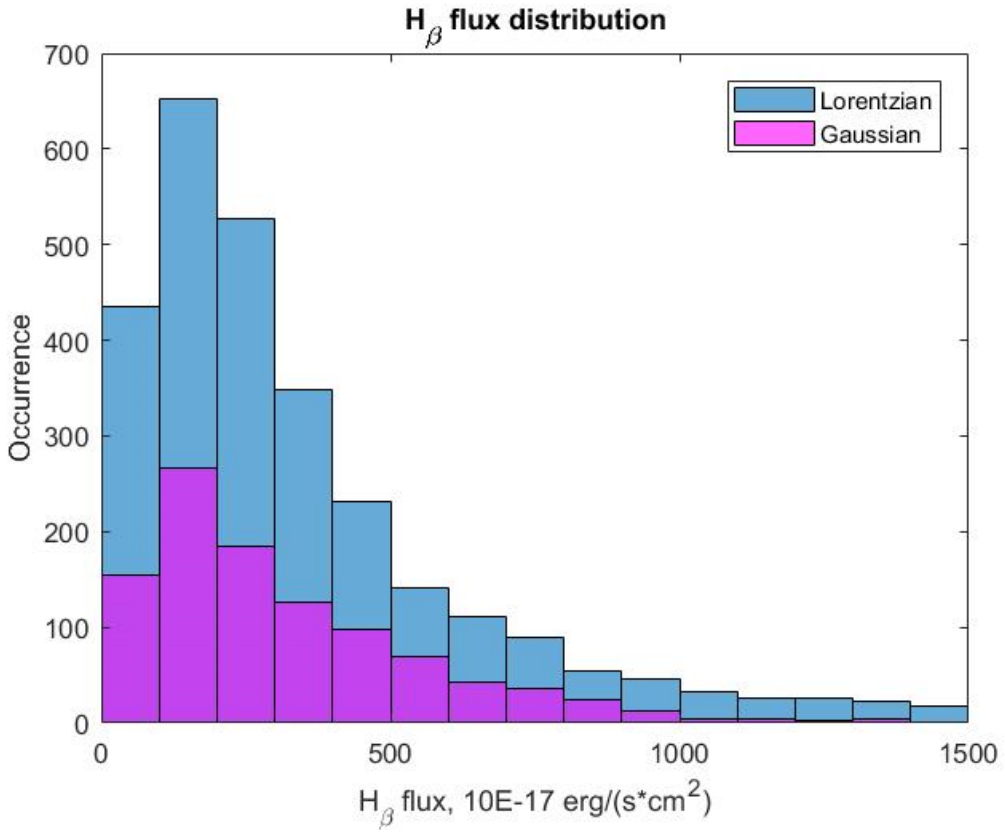

Figure 11.3 Distribution of the broad $H\beta$ flux component. The unit of the $H\beta$ flux in this histogram is $\frac{erg}{s*cm^2}$. The values in the histogram are a multiple of 10^{-17}. $H\beta$ flux is located on the x-axis, while the occurrence is on the y-axis. The Lorentzian profiles are in blue while Gaussian profiles are in magenta.

The Hβ flux distribution, seen in Figure 11.3, looks similar for both profile types, resembling a Rayleigh distribution. The most common Hβ flux values for both profiles are between 100 and 200 $10^{-17}\frac{\text{erg}}{\text{s}*\text{cm}^2}$. Though the distributions look relatively similar, the p-value obtained from the KS- test is $1.68 * 10^{-6}$, meaning that the *null* hypothesis is rejected for this parameter too. The Gaussian mean is 323.7869, the median is 254, and the standard deviation is 572.4063. For the Lorentzian profiles, the mean is 529.6636, the median is 272, and the standard deviation is 918.3652.

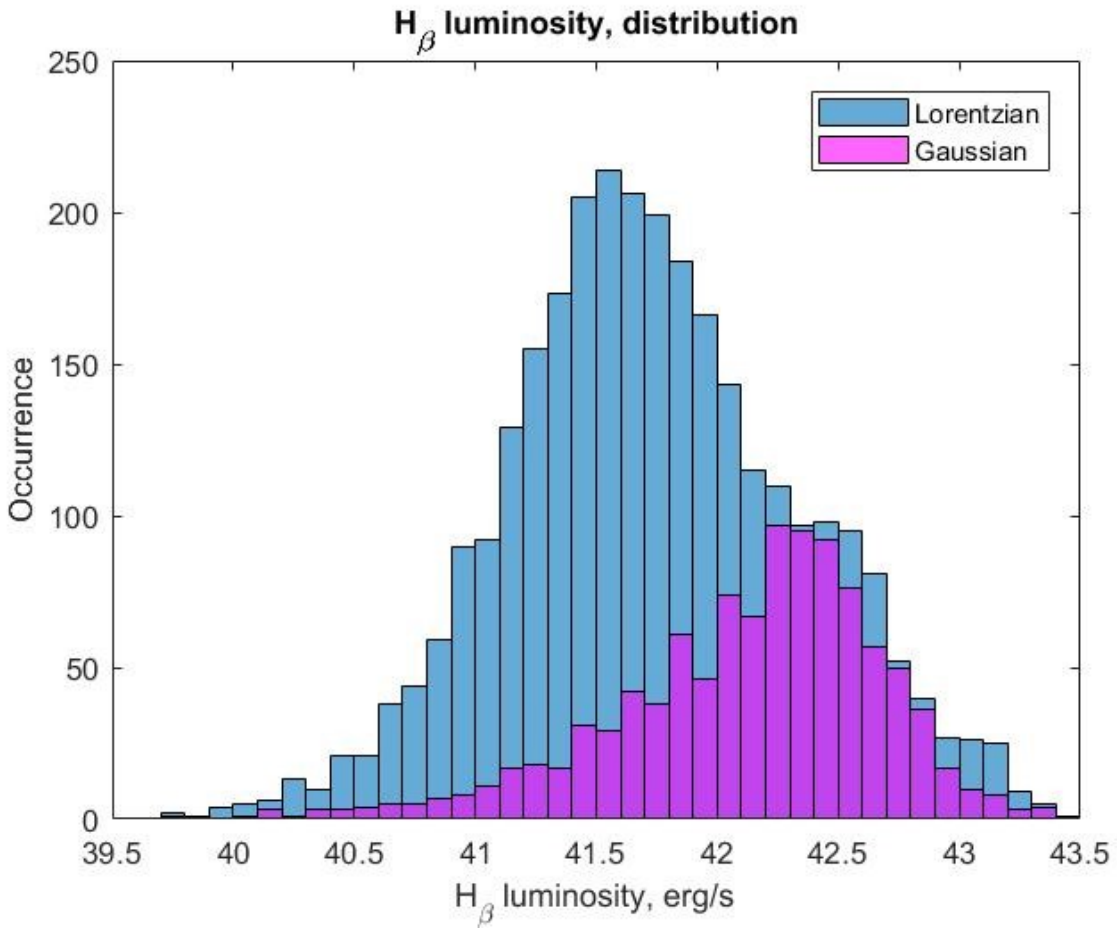

Figure 11.4 Distribution of the $H\beta$ luminosity. The unit of the $H\beta$ luminosity in this histogram is $\frac{erg}{s}$. $H\beta$ luminosity is on the x-axis, while the occurrence is on the y-axis. The Lorentzian profiles are in blue and the Gaussian profiles are in magenta.

Figure 11.4 shows the distribution of the Hβ luminosity. The Lorentzian profiles have a distribution that resembles the standard normal distribution for the Hβ luminosity. The right side of the plot, however, does not hold fully true to the standard normal

distribution. This is similar to the Gaussian redshift distribution, as in that one, the right side of the plot was also relatively significantly different from the left side of the plot. The Gaussian profiles for the $\mathrm{H}\beta$ luminosity distribution, show a shape similar to Weibull distribution. The most common $\mathrm{H}\beta$ luminosity values for the Lorentzian profiles are between 41.4 and 41.6. The Gaussian maximum values are between 42.2 and 42.5. The Gaussian mean value is 42.1275, the median is 42.2312, and the standard deviation is 0.5507. The Lorentzian values are 41.7326 for the mean, 41.6984 for the median, and 0.6085 for the standard deviation. For these two distributions, it is easy to see that the KS- test gave a p-value that rejected the *null* hypothesis. The p-value is $4.13 * 10^{-84}$.

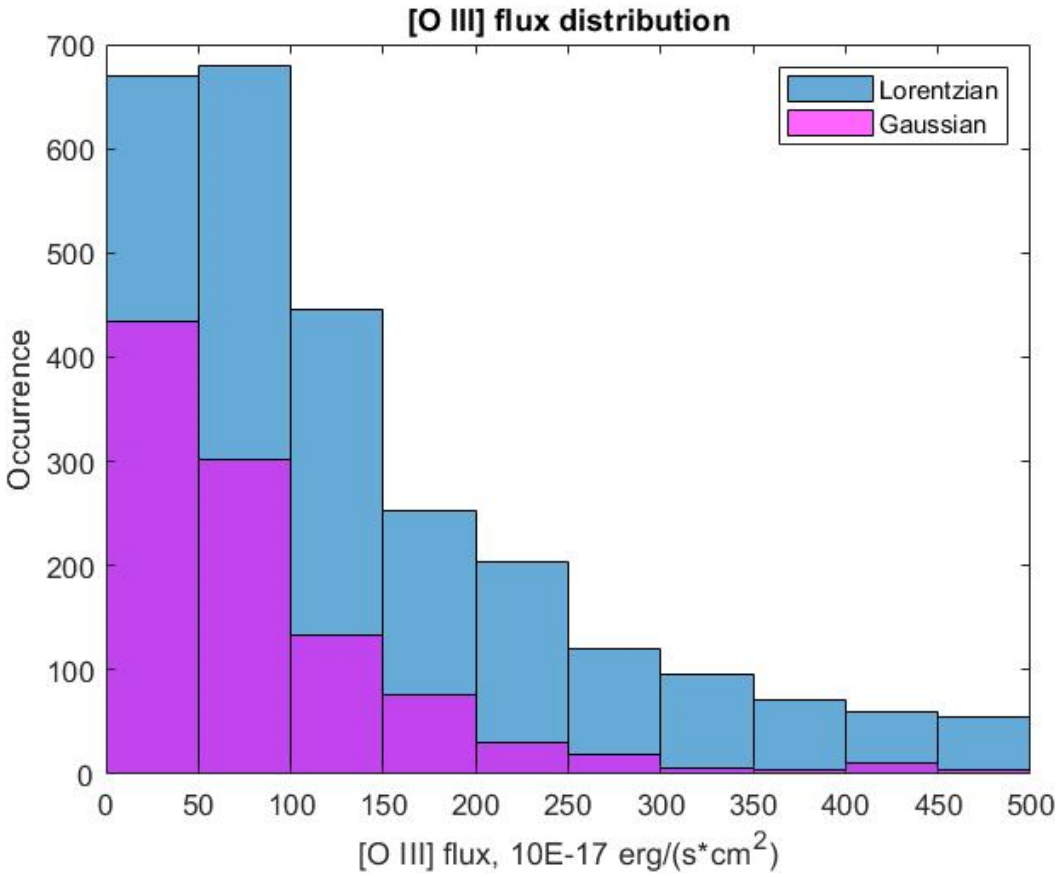

Figure 11.5 Distribution of the $[O\ III]$ flux. The unit of the $[O\ III]$ flux in this histogram is $\frac{erg}{s*cm^2}$. The values in the histogram are a multiple of 10^{-17}. $[O\ III]$ flux is on the x-axis, while the occurrence is on the y-axis. The Lorentzian profiles are in blue and the Gaussian profiles are in magenta.

The distribution for the [O III] flux, as seen in Figure 11.5, is a Poisson distribution for both profile types. The most common interval of the [O III] flux for the Gaussian profiles is the first interval, 0 to 50. For the Lorentzian profiles the most common interval is narrowly the second one, 50 to 100. The second most common interval for the Lorentzian profiles is the first one. The mean for the Gaussian profiles is 94.3857, the median is 60, and the standard deviation is 125.5139. The Lorentzian profiles have a mean of 237.6856, median of 113, and a standard deviation of 441.13114. The p-value from the KS- test is $1.54 * 10^{-46}$, which means that the *null* hypothesis is once again rejected.

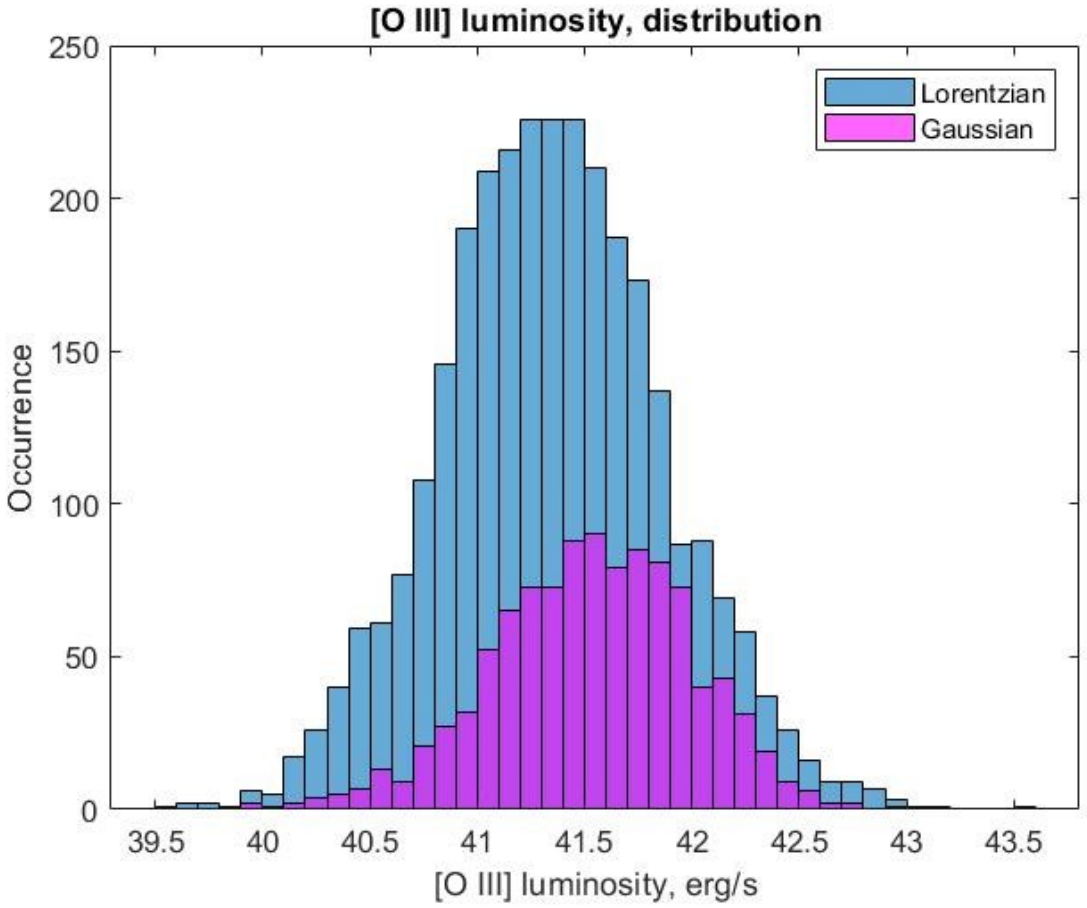

Figure 11.6 Distribution of the [O III] luminosity. The unit of the [O III] luminosity in this histogram is $\frac{erg}{s}$. [O III] luminosity is on the x-axis, while the occurrence is on the y-axis. The Lorentzian profiles are in blue and the Gaussian profiles are in magenta.

[O III] luminosity distribution is shown in Figure 11.6. The values in this histogram are similar to that of the values from Figure 11.4, Hβ luminosity histogram. The Lorentzian profiles have a distribution pattern similar to that of the leptokurtic distribution. The leptokurtic distribution is signified by the rapid increase and the wider top, though the entire distribution itself remains relatively narrow. The Gaussian profiles have a distribution more like the mesokurtic distribution. The mesokurtic distribution has a slower increase to the maximum values, but the top is still wide compared to the width of the distribution. Most common Gaussian values are between 41.4 and 41.6, the respective values for the Lorentzian profile is 41.2 and 41.5. Gaussian profiles have a mean of 41.5362, median of 41.5481, and a standard deviation of 0.4677. The mean of the Lorentzian profiles is 41.3483, the median is 41.3361, and the standard deviation is 0.5194. The p-value for the [O III] luminosity is $1.52 * 10^{-23}$, which means that the *null* hypothesis is once again rejected.

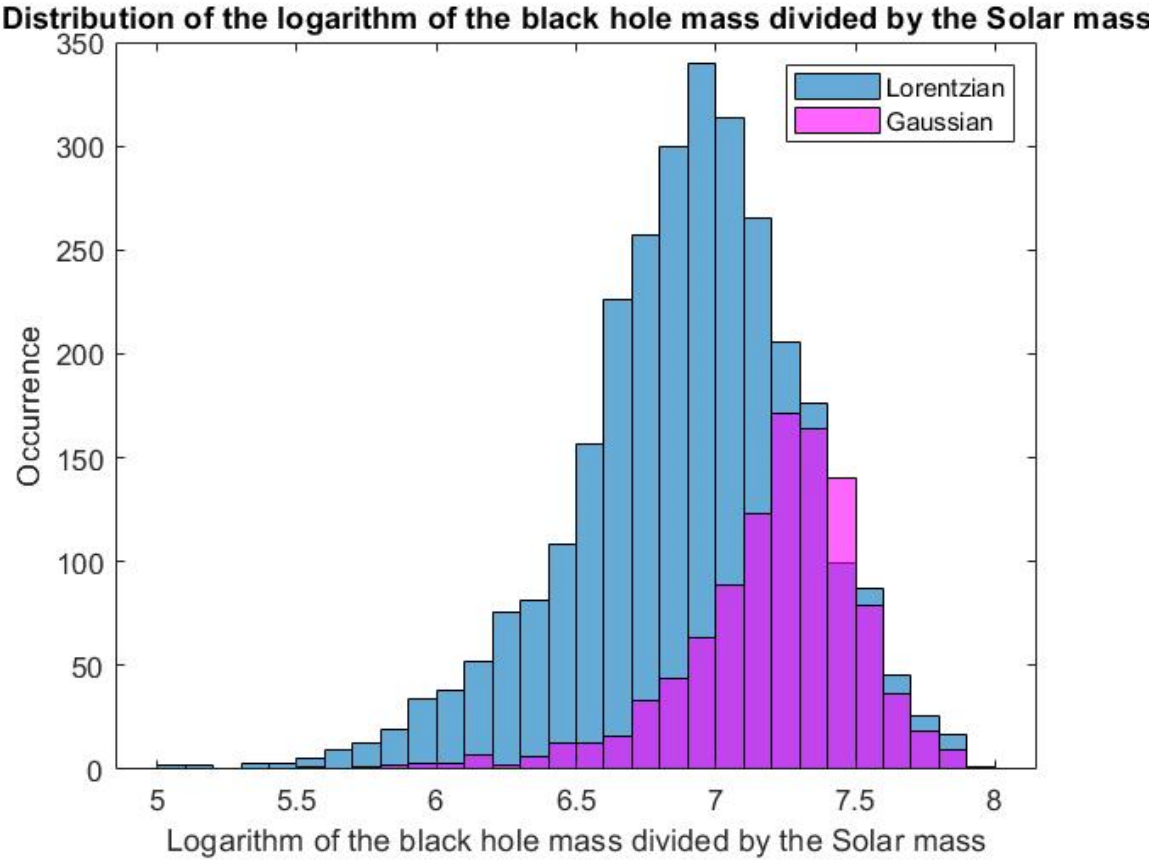

Figure 11.7 Distribution of the logarithm of the quotient of the black hole mass and the Solar mass. The logarithm of the quotient is marked on the x-axis while the occurrence is marked on the y-axis. The Gaussian sources are marked with magenta, while the Lorentzian sources are marked with blue.

Figure 11.7 shows the distribution for the logarithm of the black hole mass divided by the Solar mass. The distribution for both of the profile types resembles left-skewed with the mean for both (Gaussian: 7.2073 and Lorentzian:6.8955) being on the left side of the median (Gaussian: 7.2592 and Lorentzian: 6.9277). The standard deviation for the profile types are 0.32 for Gaussian and 0.4125 for Lorentzian. A clear distribution type is not recognized. The KS- test for the dataset of logarithm of the quotient of the black hole mass and the Solar mass returns a p-value of $2.12 * 10^{-115}$, thus rejecting the *null* hypothesis. The most common values for the Lorentzian profiles are between 6.8 and 7.1, while the respective values for the Gaussian profiles are between 7.2 and 7.4.

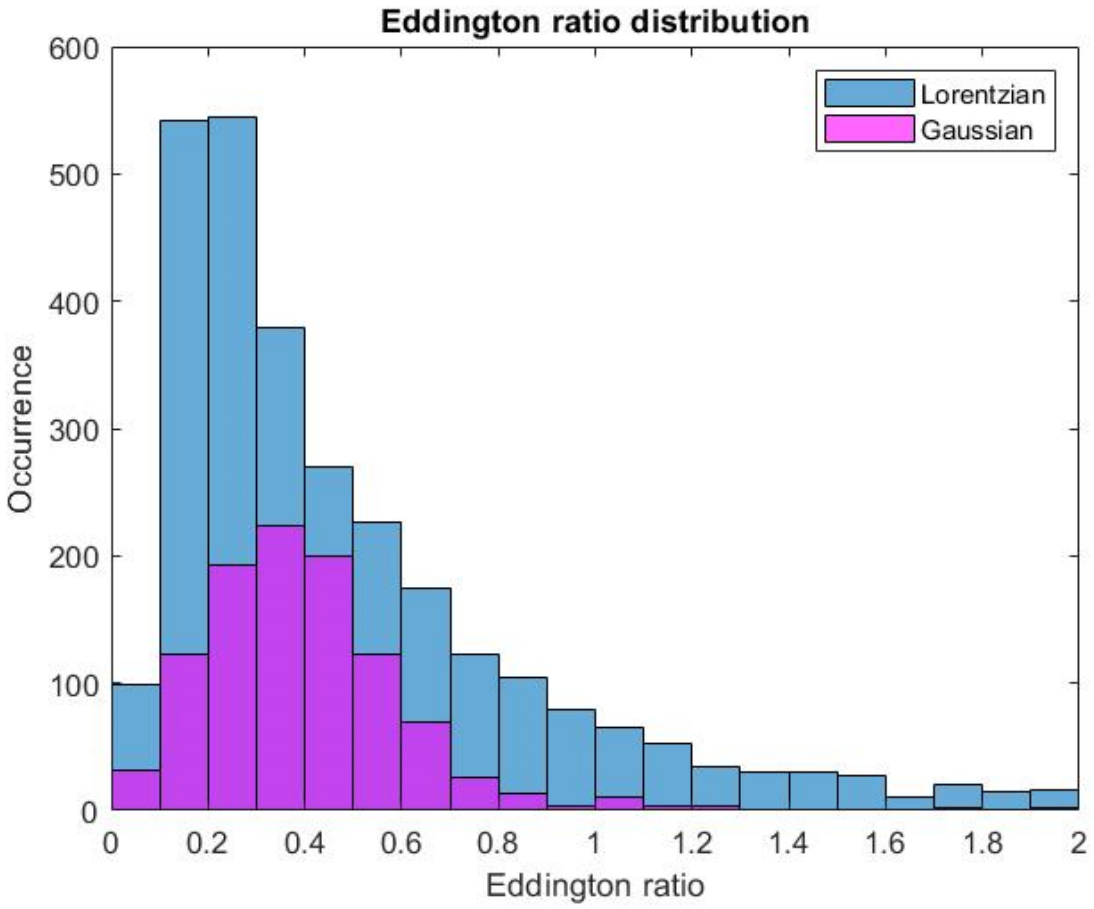

Figure 11.8 The Eddington ratio distribution, with the Eddington ratio on the x-axis and the occurrence on the y-axis. In this figure, the Lorentzian sources are marked with blue and the Gaussian sources are marked with magenta.

Figure 11.8 shows the Eddington ratio distribution. The Gaussian profiles have a binomial distribution, located on the right side of the graph. The Lorentzian profiles have a distribution similar to a logarithmic distribution. These distributions are very easy to see to be different, thus resulting in a rejected *null* hypothesis for the KS- test. The p-value of the KS- test is $5.05 * 10^{-21}$. The most common Eddington ratio for the Gaussian profiles are between 0.2 and 0.5. For the Lorentzian profiles, the most common values are between 0.1 and 0.3. The mean, median, and standard deviation values for the Gaussian profiles are 0.4257, 0.3751, and 0.4139 in order. The respective values for the Lorentzian profiles are 0.6012, 0.3786, and 0.7503.

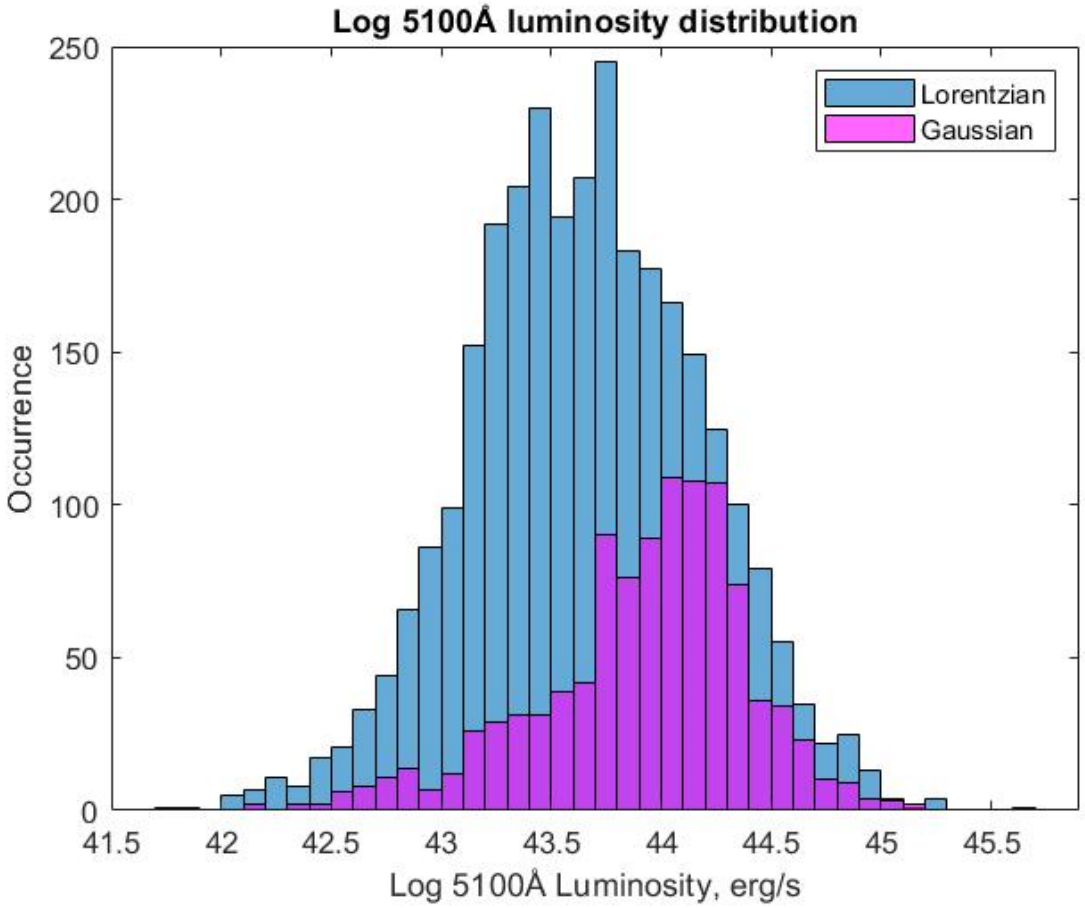

Figure 11.9 Distribution of the logarithmic nuclear monochromatic luminosity at 5100 Å. The unit for the luminosity is $\frac{erg}{s}$ and it is located on the x-axis, with the occurrence on the y-axis. The Gaussian profiles are in magenta and the Lorentzian profiles are in blue.

The logarithm of the 5100 Å luminosity is seen in Figure 11.9. Like in Figure 11.8, Figure 11.9 also shows two different type of distribution profiles for the Lorentzian and Gaussian sources. The Lorentzian sources have a leptokurtic-like distribution profile, while the Gaussian sources a more binomial distribution like distribution. The mean for the Gaussian profiles is 43.9178, the median is 44, and the standard deviation is 0.4824. The most common Gaussian values are between 44 and 44.3. For the Lorentzian profiles, the mean is 43.6583, the median is 43.65, and the standard deviation is 0.5277. The most common values for the Lorentzian profiles are lower than the values for Gaussian profiles as well as narrower. The most common Lorentzian value is 43.7, with the 43.4 being the second most common value. The p-value from the KS- test is again under 0.01, $8.27 * 10^{-53}$, so the *null* hypothesis is rejected.

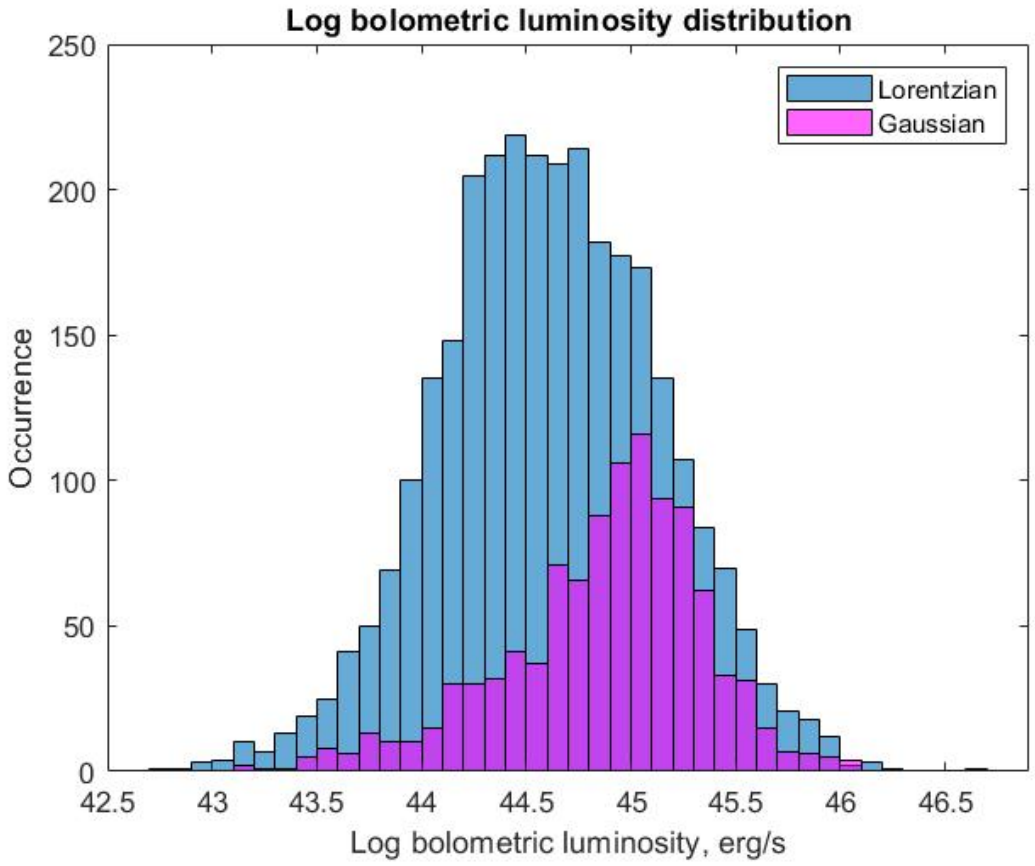

Figure 11.10 Distribution of the logarithmic bolometric luminosity measured in $\frac{erg}{s}$. The luminosity values are portrayed on the x-axis, while the occurrence is on the y-axis. Lorentzian sources are in blue, while the Gaussian sources are marked with magenta.

Figure 11.10 shows the distribution of the logarithm of the bolometric luminosity. The distributions of Figure 11.10 are very similar in shape to the distributions in Figure 11.9 for the respective profile types. From this, it is easy to decipher that the KS- test once again returns a rejected *null* hypothesis with a p-value of $8.27 * 10^{-53}$. The mean, median, and standard deviation values for the Gaussian profiles are 44.872, 44.9542, and 0.4824 in order. The first two values are a little higher than the respective values for the logarithmic nuclear monochromatic luminosity at 5100 Å portrayed in Figure 11.9. The standard deviation remains the same for the Gaussian profile types in both Figure 11.10 and Figure 11.9. The mean, median, and standard deviation for the Lorentzian profiles are 44.6125, 44.6042, and 0.5277. The standard deviation for the Lorentzian profiles is the same for both Figure 11.9 and Figure 11.10. The other two values (mean and median) are, however, larger for Figure 11.10 than for Figure 11.9. The most common Gaussian values are between 44.9 and 45.0. For the Lorentzian profiles, the most common values are between 44.2 and 44.8.

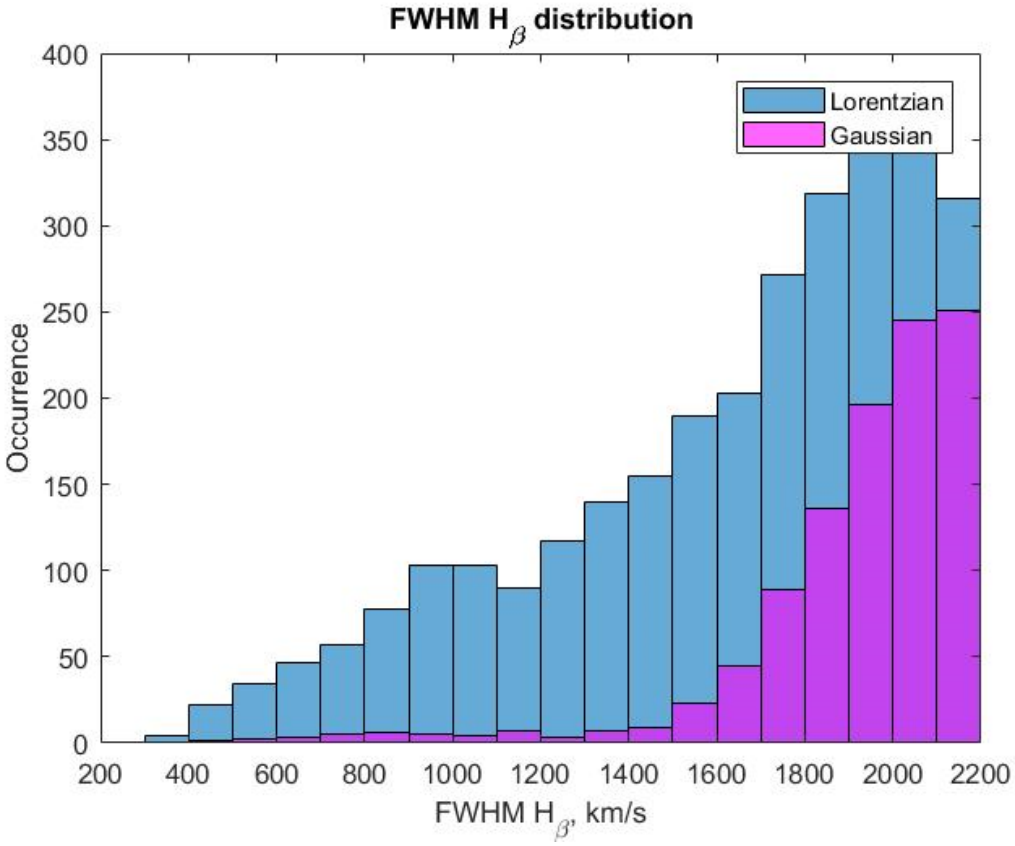

Figure 11.11 Distribution of the broad $H\beta$ component FWHM. The unit for FWHM is $\frac{km}{s}$ and the variable is located on the x-axis. The occurrence can be found on the y-axis. Blue bars represent the Lorentzian profiles, while the magenta bars represent Gaussian profiles.

Distribution of the FWHM Hβ can be seen in Figure 11.11. The unit of FWHM Hβ used is km/s, meaning it is not from the CGS unit system. The reason why km/s was used instead of cm/s is because the km/s is generally believed to be easier to understand size-wise than cm/s. The Lorentzian profiles have a steadier, slower growth than the Gaussian profiles which have a very rapid increase at the very right side of the graph. The Gaussian profiles have a mean of 1927.1, median of 1989, and standard deviation of 259.489. For the Lorentzian profiles, these values are 1623.28, 1752, and 433.405. The p-value obtained from the KS- test is $6.46 * 10^{-81}$, meaning that the *null* hypothesis is rejected. The most common Lorentzian FWHM values are between 1900 and 2100. The respective Gaussian values are 2000 and 2200.

Table 10.1 Short summary of the results

	Lorentzian	**Gaussian**
Redshift	A significant number of sources had a low redshift value, most sources had a redshift between 0.19-0.22.	More evenly spread out distribution, relatively many sources had a much higher redshift value than most Lorentzian profile types.
R4570	Relatively similar distribution for both fit types most sources had R4570 values between 0.5 and 0.7	Most gaussian sources had a R4570 value between 0.5 and 0.7.
Hβ flux	Similar results again. Most sources had values between 100 and 200.	Most common Hβ flux values between 100 and 200.

$H\beta$ luminosity	Significantly more likely to have lower luminosity values than Gaussian profiles. Most common luminosity values between 41.4 and 41.6.	Much more likely to have higher luminosity values than Lorentzian profiles. Most common luminosity values from 42.2 to 42.5.
$[O\ III]$ flux	Similar distribution, but a wider top for Lorentzian. The most common values between 0 and 100.	Most common values between 0 and 50.
$[O\ III]$ luminosity	Slightly smaller luminosity values on average. Most common values between 41.2 and 41.5	Most common luminosity values between 41.4 and 41.6, so the difference between the profiles is very small
Logarithm of the quotient of black hole mass divided by Solar mass	On average, the black hole mass is significantly smaller for the Lorentzian profile sources. Most common mass between 6.8 and 7.1.	Significantly heavier black hole masses. Most common black hole mass between 7.2 and 7.4.
Eddington ratio	More variation in the Eddington ratio, slowly decreasing popularity of values starting from 0.1. Most common values between 0.1 and 0.3.	The distribution is very clumped up between 0.1 and 0.7. Most common values between 0.3 and 0.4.

Log 5100 Å luminosity	Usually slightly lower values. Most common values 43.7 and 43.4.	Significantly more common to have slightly higher values. Most common values between 44 and 44.3.
Log bolometric luminosity	Very wide range of most common values (44.2 to 44.8). Lower luminosity values more common than for Gaussian profiles.	Most common values between 44.9 and 45.0. A lot of sources with higher luminosity value than for Lorentzian profiles.
FWHM $H\beta$	More even growth of FWHM values. Most common values between 1900 and 2100.	A relatively sudden growth to the most common FWHM values (2000 to 2200).

12 Discussion

This chapter highlights the biggest differences in the analyzed parameters seen in Chapter 11. The properties of the source galaxy can vary depending on which profile type (Lorentzian or Gaussian) fits the source better. The differences in the properties vary in scale, as parameters sometimes only had very small differences while sometimes the difference between the parameters was relatively large. In the used sample, it appears that the sources with a Gaussian fit tend to on average be slightly brighter than the Lorentzian fitted sources are on average. Sources with a better Gaussian profile type have, without an exception in this thesis, a higher average luminosity value for all different luminosity types (logarithmic 5100 Å luminosity, Hβ luminosity, [O III] luminosity, and bolometric luminosity). The most major difference between the Lorentzian and Gaussian profile types for a luminosity parameter can be observed in Figure 11.4, Hβ luminosity. The closest or most similar luminosity values for Lorentzian and Gaussian profiles occur for the [O III] luminosity parameter.

When observing the redshift distribution from Figure 11.1, most of the Lorentzian sources have a significantly lower redshift value compared to the redshift value of many of the Gaussian sources. From Figure 11.1, it appears like roughly half of the Gaussian fitted sources have a higher redshift than most Lorentzian fitted sources. However, Lorentzian profiled sources could also have a high redshift value (above 0.5), but most the sources had a lower (below 0.5) redshift value. Gaussian profiled sources have relatively evenly distributed redshift values. A higher redshift value means that the source is located further away than a source with a lower redshift value. This means that most Lorentzian profiled NLS1 galaxies in the used data sample are located closer to us than many of the Gaussian profiled NLS1 galaxies from this data sample.

Figure 11.7 shows the distribution of the logarithmic black hole mass divided by the Solar mass. The shape of the distribution is similar for both the Lorentzian fitted sources and the Gaussian fitted sources. The values for the Gaussian profiles are, however, centered at higher values (7.2 to 7.4) than the Lorentzian fitted sources (6.9 to 7.0). The spread of the mass values is much wider for the Lorentzian fitted sources than for the Gaussian fitted sources. Though it is more common for a Gaussian fitted source to have higher mass values, also Lorentzian sources can have similarly high mass values.

The Eddington ratio, seen in Figure 11.8, is distributed relatively differently between the two profile types. While the most common Eddington ratio values are very similar to both profiles, the shape of the distribution is very different. Lorentzian sources have a wider spread of values and decreasing popularity of values after the most common values (0.1 to 0.3). Gaussian sources have increasing popularity of values from 0 to 0.4. After

0.4, the values have decreasing popularity. The Eddington ratio for the Gaussian sources is mostly concentrated between 0.2 and 0.5.

The last parameter observed in this thesis to show a significant difference in the distribution for the two profile types is the broad $H\beta$ component FWHM seen in Figure 11.11. The most common values for the profile types are both located on the far-right side of the graph, so on the faster side of the graph. The Lorentzian fitted sources show once again a wider spread of common values. Slower FWHM values are far more likely for the Lorentzian fitted sources as there are only very few slow Gaussian sources. The rest of the parameters observed in this thesis (R4570, $H\beta$ flux, and $[O\ III]$ flux) have a relatively similar in both shape and for values for the two profile types.

13 Conclusions

The very main finding in this thesis is that of the 11001 sources Rakshit et al. (2017) found to be NLS1 galaxies, only 3998 can be confirmed in this thesis to be with high likelihood NLS1 galaxies. Of the remaining 7001 galaxies, two sources were discarded due to issues with the code. Due to the very strict criteria used, NLS1 galaxies were lost in the discarding process. The two sources that were omitted due to issues with the code could not be verified to either not be or to be NLS1 galaxies with a high likelihood. To complete further analysis on the 3998 sources and their parameters, these sources were divided into two categories. These categories are based on what profile type fit the source better: Gaussian or Lorentzian. The parameters for both of these profile types were then analyzed in Chapter 11 with methods described in Chapter 10. Of the 3998 sources, 1037 sources had a better Gaussian fit and the remaining 2961 sources had a better Lorentzian fit. This means that the Lorentzian profile fit was in this dataset far more likely than the Gaussian profile fit. Sources with a better Lorentzian fit are viewed to be very likely NLS1 galaxies. Gaussian fitted sources are deemed quite likely NLS1.

The parameters observed in thesis can be divided into two categories: parameters with significant difference and parameters without significant difference. The parameters that showed significant differences in distribution and/or values for the two profile types are all the luminosity parameters, redshift, black hole mass, Eddington ratio, and FWHM. The parameters that did not have significant differences in either the distribution or the values are both the flux parameters, R4570, and $\frac{\mathrm{H}\beta}{M_{\mathrm{BH}}}$. As according to the obtained results on the used data set, the profile type seems to be indifferent for the values and distribution shape of the parameters with no significant differences. These parameters could thus be viewed as nearly redundant for future work on determining differences between the profile types for the NLS1 galaxies.

As noted in Chapter 12, the Gaussian fitted sources had a higher average luminosity value for all the different luminosity parameters than the average luminosity value of the Lorentzian sources. This appear to mean that if a dimmer NLS1 galaxy is observed, the profile type for this source is more likely to be Lorentzian fitted than Gaussian fitted. Since bright Lorentzian sources do also exist, a bright NLS1 galaxy could have either one of the two profile types.

The redshift of the Lorentzian fitted sources is on average on the lower end (0.4 and under) of the graph. The Gaussian fitted sources have a very even distribution of redshift values, with both ends of the redshift spectrum appearing relatively regularly. While the

Lorentzian sources are not very likely to have high redshift values (0.6 and over), Gaussian sources have these higher redshifts relatively frequently. All of this means that if an NLS1 galaxy has a low redshift value (i.e. 0.12) it is with very high likelihood that the source has a Lorentzian profile type. The very low redshift values (between 0 and 0.4) are, according to the data in this thesis, in nearly all cases Lorentzian fitted. This means that the sources close to us, are in almost all cases, Lorentzian fitted. The sources that are located further away from us can be either Lorentzian or Gaussian, but sources that are located extremely far from us (redshift over 0.62) are very likely to be Gaussian sources.

Low logarithm values of the quotient of the black hole mass divided by the Solar mass are more common for Lorentzian fitted sources than for Gaussian fitted sources. The Lorentzian sources have a wider spread of values compared to the Gaussian values. On average, Gaussian sources tend to have a high black hole mass value, with very few low black hole mass values. When an NLS1 galaxy is observed and it has a low black hole mass, according to the obtained data it is exceedingly likely that the galaxy has a Lorentzian profile type. Both profile types have sources with a high black hole mass, so a new source with a high black hole mass could be of either profile type. There are very similar amounts of the high black hole massed sources for the two profile types. Since this dataset has a lot more of Lorentzian fitted sources than Gaussian fitted sources, it is possible that the sources with heavier black hole masses would actually be more common for the Gaussian fits than for the Lorentzian fits. This is, however, just an assumption.

The Eddington ratio for Gaussian fitted sources are very concentrated between 0.1 and 0.7, with basically no sources with higher Eddington ratios. Lorentzian fitted sources have both high and low Eddington ratios, with the low Eddington ratios (between 0.1 and 0.3) being far more common. From this, it is possible to draw the conclusion that sources with a highish Eddington ratio (above 0.7) are very likely to be Lorentzian fitted sources.

With the results of this thesis and a cleaned data sample, it is possible to start analyzing clean NLS1 sources, and not have the results contaminated due to non-NLS1 sources. Studying the differences between the Lorentzian and Gaussian profiled NLS1 sources in this thesis, could be used in the future to further understand why there are such differences between the profile types. The results of this thesis will be used for a refereed article.

References

Alonso, S., Mesa, V., Padilla, N. and Lambas, D. (2012). Galaxy interactions. Astronomy & Astrophysics, [online] 539, p.A46. Available at: https://www.aanda.org/articles/aa/abs/2012/03/aa17901-11/aa17901-11.html [Accessed 7 Mar. 2019].

Bouton, E. (2014). Finding Aid to the Papers of Marshall H. Cohen 1957-2002. (online) Nrao.edu. Available at: https://www.nrao.edu/archives/Cohen/cohen.shtml (Accessed 18 Nov. 2018).

Brunthaler, A., Falcke, H., Bower, G., Aller, M., Aller, H., Teräsranta, H., Lobanov, A., Krichbaum, T. and Patnaik, A. (2000). III Zw 2, the first superluminal jet in a Seyfert galaxy. Astronomy and Astrophysics, (online) 357, p45-48. Available at: http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?2000A%26A...357L..45B&data_type=PDF_HIGH&whole_paper=YES&type=PRINTER&filetype=.pdf (Accessed 18 Nov. 2018).

Byrd, D. (2017). Today in science: Quasar mystery solved | EarthSky.org. (online) Earthsky.org. Available at: https://earthsky.org/space/this-date-in-science-maartin-schmidt-discovers-first-known-quasar (Accessed 18 Nov. 2018).

Carilli, C. and Harris, D. (1996). Cygnus A. Cambridge: Cambridge University Press 264.

Chen, Y. and Hwang, C. (2017). Morphology of Seyfert galaxies. Astrophysics and Space Science, (online) 362(12). Available at: https://arxiv.org/pdf/1707.08806.pdf (Accessed 20 Nov. 2018).

Combes, F. (2006). Secular evolution in galaxies. Proceedings of the International Astronomical Union, [online] 2(S235), p.19. Available at: http://adsabs.harvard.edu/abs/2007IAUS..235...19C [Accessed 7 Mar. 2019].

Decarli, R., Dotti, M., Fontana, M. and Haardt, F. (2008). Are the black hole masses in narrow-line Seyfert 1 galaxies actually small?. Monthly Notices of the Royal Astronomical Society: Letters, (online) 386(1), pL15-L19. Available at: http://adsabs.harvard.edu/abs/1985ApJ...297..166O (Accessed 28 Jan. 2019).

Dibaj, E. (1981). On the difference between Seyfert 1 and Seyfert 2 galaxies. (online) Adsabs.harvard.edu. Available at: http://adsabs.harvard.edu/abs/1981PAZh....7..451D (Accessed 19 Nov. 2018).

En.wikiquote.org. (n.d.). Talk:Socrates - Wikiquote. (online) Available at: https://en.wikiquote.org/wiki/Talk:Socrates (Accessed 15 Oct. 2018).

Galaxy Zoo (2010). Hubble's galaxy classification scheme. (image) Available at: http://zooniverse-resources.s3.amazonaws.com/blogs.zooniverse.org/5/files/2010/05/HubbleTuningFork2w.jpg (Accessed 16 Oct. 2018).

Ghisellini, G., Tavecchio, F., Maraschi, L., Celotti, A. and Sbarrato, T. (2014). The power of relativistic jets is larger than the luminosity of their accretion disks. Nature, (online) 515(7527), p376-378. Available at: https://www.nature.com/articles/nature13856 (Accessed 10 Dec. 2018).

Goad, M., Korista, K. and Ruff, A. (2012). The broad emission-line region: the confluence of the outer accretion disc with the inner edge of the dusty torus. Monthly Notices of the Royal Astronomical Society, (online) 426(4), p3086-3111. Available at: https://arxiv.org/abs/1207.6339 (Accessed 28 Jan. 2019).

Järvelä, E. (2018). Narrow-line Seyfert 1 galaxies - Observational and statistical analysis. Doctoral dissertations. Aalto University.

Karttunen, H., Kröger,, Oja, H., Poutanen, M. and Donner, K. (2016). Fundamental Astronomy. 6th ed. Helsinki: Springer.

Komossa, S. (2018). Multi-Wavelength Properties of Radio-loud Narrow-Line Seyfert 1 Galaxies. Proceedings of Science. [online] Available at: https://arxiv.org/pdf/1807.03666.pdf [Accessed 5 Mar. 2019].

Kormendy, J. (2012). Internal and environmental secular evolution of disk galaxies. Proceedings of the International Astronomical Union, [online] 10(H16), pp.316-317. Available at: http://adsabs.harvard.edu/abs/2015HiA....16..316K [Accessed 7 Mar. 2019].

Kormendy, J. (2013). Secular Evolution in Disk Galaxies. Cambridge University Press. [online] Available at: https://arxiv.org/abs/1311.2609v1 [Accessed 7 Mar. 2019].

Koulouridis, E., Plionis, M., Chavushyan, V., Dultzin, D., Krongold, Y., Georgantopoulos, I. and León-Tavares, J. (2013). Activity of the Seyfert galaxy neighbours. Astronomy & Astrophysics, (online) 552, A135. Available at: https://www.researchgate.net/publication/258843511_Activity_of_the_Seyfert_galaxy_neighbours_Koulouridis_2013 (Accessed 28 Jan. 2019).

La Mura, G., Busetto, G., Ciroi, S., Rafanelli, P., Berton, M., Congiu, E., Cracco, V. and Frezzato, M. (2017). Relativistic plasmas in AGN jets. The European Physical Journal D, [online] 71(4). Available at: https://arxiv.org/abs/1702.06779 [Accessed 12 Apr. 2019].

Lähteenmäki, A. and Valtaoja, E. (1999). Optical Polarization and Imaging of Hot Spots in Radio Galaxies. The Astronomical Journal, [online] 117(3), pp.1168-1174. Available at: https://iopscience.iop.org/article/10.1086/300752/fulltext/970374.text.html [Accessed 5 Mar. 2019].

Lähteenmäki, A., Järvelä, E., Hovatta, T., Tornikoski, M., Harrison, D., López-Caniego, M., Max-Moerbeck, W., Mingaliev, M., Pearson, T., Ramakrishnan, V., Readhead, A., Reeves, R., Richards, J., Sotnikova, Y. and Tammi, J. (2017). 37 GHz observations of narrow-line Seyfert 1 galaxies. Astronomy & Astrophysics, [online] 603, p.A100. Available at: https://aaltodoc.aalto.fi/handle/123456789/28143 [Accessed 12 Apr. 2019].

Lähteenmäki, A., Järvelä, E., Ramakrishnan, V., Tornikoski, M., Tammi, J., Vera, R. and Chamani, W. (2018). Radio jets and gamma-ray emission in radio-silent narrow-line Seyfert 1 galaxies. Astronomy & Astrophysics, [online] 614, p.L1. Available at: https://aaltodoc.aalto.fi/handle/123456789/32711 [Accessed 12 Apr. 2019].

Lemke, D., Stickel, M. and Wilke, K. (2000). ISO Surveys of a Dusty Universe. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Lyu, J. and Rieke, G. (2017). The Intrinsic Far-infrared Continua of Type-1 Quasars. The Astrophysical Journal, [online] 841(2), p.76. Available at: https://arxiv.org/abs/1704.06987 [Accessed 12 Apr. 2019].

Maddox, S., Sutherland, W., Efstathiou, G., Loveday, J. and Peterson, B. (1990). Galaxy Evolution at Low Redshift. Monthly Notices of the Royal Astronomical Society, [online] 247(1). Available at: http://adsabs.harvard.edu/abs/1990MNRAS.247P...1M [Accessed 7 Mar. 2019].

Maia, M., Machado, R. and Willmer, C. (2003). Are Seyfert types 1 and 2 located in different morphological and luminosity hosts?. Bulletin of the Astronomical Society of Brazil, (online) 1. Available at: http://Research Gate (Accessed 20 Nov. 2018).

Malkan, M. and Sargent, W. (1982). The ultraviolet excess of Seyfert 1 galaxies and quasars. The Astrophysical Journal, (online) 254 22. Available at: http://articles.adsabs.harvard.edu/cgi-bin/nph-

iarticle_query?1982ApJ...254...22M&data_type=PDF_HIGH&whole_paper=YES&type=PRINTER&filetype=.pdf (Accessed 20 Nov. 2018).

Marshall, J. and Marshall, V. (2008). Rediscovery of the Elements: Mineral Waters and Spectroscopy. The Hexagon, (online) 99(3), p42-46. Available at: https://chemistry.unt.edu/sites/default/files/users/owj0001/bunsen%20and%20spectroscopy.pdf (Accessed 10 Dec. 2018).

Massaglia, S. (2009). General Properties of Jets from Active Galactic Nuclei and Comparison with Protostellar Jets. Protostellar Jets in Context, [online] pp.273-281. Available at: https://www.researchgate.net/publication/227023938_General_Properties_of_Jets_from_Active_Galactic_Nuclei_and_Comparison_with_Protostellar_Jets [Accessed 12 Apr. 2019].

Mathur, S. (2000). Narrow-line Seyfert 1 galaxies and the evolution of galaxies and active galaxies. Monthly Notices of the Royal Astronomical Society, (online) 314(4), pL17-L20. Available at: http://adsabs.harvard.edu/abs/2000MNRAS.314L..17M (Accessed 28 Jan. 2019).

Mundell, C., Ferruit, P., Nagar, N. and Wilson, A. (2009). RADIO VARIABILITY IN SEYFERT NUCLEI. The Astrophysical Journal, [online] 703(1), pp.802-815. Available at: https://arxiv.org/abs/0907.1489v1 [Accessed 5 Mar. 2019].

NRAO/AUI/NSF (n.d.). Viewing angle comparison. (image) Available at: https://astronomy.com/-/media/Images/News%20and%20Observing/News/2018/07/BlazarCartoon.jpg?mw=1000&mh=800 (Accessed 17 Nov. 2018).

Onken, C. and Peterson, B. (2002). The Mass of the Central Black Hole in the Seyfert Galaxy NGC 3783. The Astrophysical Journal, [online] 572(2), pp.746-752. Available at: https://arxiv.org/abs/astro-ph/0202382 [Accessed 5 Mar. 2019].

Osterbrock, D. and Pogge, R. (1985). The spectra of narrow-line Seyfert 1 galaxies. The Astrophysical Journal, (online) 297, 166. Available at: http://adsabs.harvard.edu/abs/1985ApJ...297..166O (Accessed 28 Jan. 2019).

Rakshit, S., Stalin, C., Chand, H. and Zhang, X. (2017). A Catalog of Narrow Line Seyfert 1 Galaxies from the Sloan Digital Sky Survey Data Release 12. The Astrophysical Journal Supplement Series, (online) 229(2), 39. Available at: https://iopscience.ioorg/article/10.3847/1538-4365/aa6971/pdf (Accessed 13 Feb. 2019).

Robson, I. (1996). Active galactic nuclei. Chichester: Wiley.

Saikia, D. (2018). Jets in Active Galaxies. Resonance, [online] 23(2), pp.147-163. Available at: https://www.ias.ac.in/article/fulltext/reso/023/02/0147-0163 [Accessed 12 Apr. 2019].

Schmidt, E., Ferreiro, D., Vega Neme, L. and Oio, G. (2016). Spectral nuclear properties of NLS1 galaxies. Astronomy & Astrophysics, (online) 596, A95. Available at: https://arxiv.org/abs/1608.02396 (Accessed 28 Jan. 2019).

Sellwood, J. (2014). Secular evolution in disk galaxies. Reviews of Modern Physics, [online] 86(1), pp.1-46. Available at: https://arxiv.org/abs/1310.0403v2 [Accessed 7 Mar. 2019].

Shrader, C. and Beckmann, V. (2013). Active galactic nuclei. Wiley-VCH.

Silk, J. and Rees, M. (1997). Quasars and Galaxy formation. (ebook) Astronomy and Astrophysics. Available at: https://arxiv.org/pdf/astro-ph/9801013.pdf (Accessed 15 Oct. 2018).

Simkin, S. and Mehlberg, L. (1989). Morphological Differences Between Type 1 and Type 2 Seyfert Galaxies. Bulletin of the American Astronomical Society, (online) 21, 776. Available at: http://adsabs.harvard.edu/full/1989BAAS...21..776S (Accessed 20 Nov. 2018).

Sparke, L. and Gallagher, J. (2007). Galaxies in the Universe. 2nd ed. Cambridge: Cambridge University Press

Urry, M. and Padovani, P. (1995). Structure of an AGN. [image] Available at: http://10.1086/133630 [Accessed 7 May 2019].

Vennes,, S., Polomski,, E., Bowyer,, S. and Thorstensen, J. (1995). Discovery of Extreme-Ultraviolet Radiation from the Seyfert Galaxy Ton S180 (=EUVE J0057−223). The Astrophysical Journal, 448(1).

Vulcani, B., Poggianti, B., Fritz, J., Fasano, G., Moretti, A., Calvi, R. and Paccagnella, A. (2014). FROM BLUE STAR-FORMING TO RED PASSIVE: GALAXIES IN TRANSITION IN DIFFERENT ENVIRONMENTS. The Astrophysical Journal, [online] 798(1), p.52. Available at: https://arxiv.org/abs/1410.6481 [Accessed 7 Mar. 2019].

Williams, D. (2018). Sun Fact Sheet. (online) Nssdc.gsfc.nasa.gov. Available at: https://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html (Accessed 16 Oct. 2018).

Yoshida, S. (2002). The Eddington Limit. (online) Ppl.phys.chiba-u.j Available at: http://www.ppl.phys.chiba-u.jp/lecture/radiation/node2.html (Accessed 28 Jan. 2019).

Zhou, Y., Yu, K., Young, E., Wang, J. and Ma, E. (1997). Statistical Properties of the Big Blue Bump in Active Galactic Nuclei. The Astrophysical Journal, (online) 475(1), pL9-L12. Available at: https://iopscience.ioorg/article/10.1086/310451 (Accessed 28 Jan. 2019).

Appendix – Marco Berton's source analysis code

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit
from scipy.stats import chisquare
c = 299792.458
res_inst = 176. #sdss in kms
res_inst = 0. #we are not interested in this
lambda0 = 4861.363
width = 50 #line width in Angstrom
res_inst = res_inst*lambda0/(2.355*c)
def lorentzian(x, flux, center, core):                                  #define lorentzian function
        gamma = 2.*flux/(np.pi*core)
        return (gamma*flux/2.)/(np.pi*((x - center)**2 + gamma**2/4.))
def gaussian(x, height, center, width):                                 #define gaussian function
        return np.fabs(height)*np.exp(-(x - np.fabs(center))**2/(2*(width**2 + res_inst**2)))
file1 = 'download_list.csv'                     #file containing plate, mjd e fiber of the objects
block = np.genfromtxt(file1, usecols=(0), dtype='|S50')
z = np.genfromtxt(file1, usecols=(1))
plate, mjd, fiber = [], [], []
for i in range(10983, 10986):
        plate = (str(block[i].split(",")[1]))
        mjd = (str(block[i].split(",")[2]))
        fiber = (str(block[i].split(",")[3]))
        if fiber == '000':
```

```
            filename = 'spec-'+plate+'-'+mjd+'-1'+fiber

        else:

            filename = 'spec-'+plate+'-'+mjd+'-0'+fiber

#################################### READING FILE AND CORRECTING FOR REDSHIFT ####################################

        wave = np.genfromtxt(str(filename)+'.txt',usecols=(0))/(1+z[i])

        flux = np.genfromtxt(str(filename)+'.txt',usecols=(1))

        x_hb = wave[np.where((wave >= lambda0 - width) & (wave <= lambda0 + width))]

        y_hb = flux[np.where((wave >= lambda0 - width) & (wave <= lambda0 + width))]

#################################### CONTINUUM SUBTRACTION ####################################

        int1 = 4700.            #interval 1 to estimate continuum, left of the lines

        int2 = 4710.

        int3 = 5050.            #interval 2 to estimate continuum, right of the lines

        int4 = 5150.

        xcont = np.concatenate([wave[np.where((wave > int1) & (wave < int2))], wave[np.where((wave > int3) & (wave < int4))]])

        ycont = np.concatenate([flux[np.where((wave > int1) & (wave < int2))], flux[np.where((wave > int3) & (wave < int4))]])

        rms_cont = np.std(ycont)

        med_cont = np.median(ycont)

        snratio = med_cont/rms_cont

#################################### FITTING THE CONTINUUM WITH A LINE ####################################

        par_cont1, par_cont2 = np.polyfit(xcont,ycont,1)

        fit_cont = []

        for j in range(0,len(x_hb)):

            fit_cont.append(par_cont1*x_hb[j] + par_cont2)
```

```
#####################################  PLOT OF THE CONTINUUM SUBTRACTION  #####################################

#       plt.plot(x_hb,fit_cont,lw=2,color='black')

#       plt.plot(x_hb, y_hb,lw=1.5,color='green')

        for j in range(0,len(y_hb)):

                y_hb[j] = y_hb[j] - fit_cont[j]

#       plt.plot(x_hb, y_hb,lw=1.5, color='red')

#       plt.show()

#####################################  LORENTZIAN FITTING WITH CURVE_FIT  #####################################

        guess_lorentz = [max(y_hb), 4861, 3]

        popt_l, pcov_l = curve_fit(lorentzian, x_hb, y_hb, p0=guess_lorentz, maxfev = 10000)

        best_lorentz = lorentzian(x_hb, popt_l[0], popt_l[1], popt_l[2])

#####################################  LORENTZIAN CHI SQUARE  #####################################

        chi_squared_lorentzian = 0

        for j in range (0,len(y_hb)):

                chi_squared_lorentzian = chi_squared_lorentzian + np.fabs((y_hb[j] - best_lorentz[j])**2/best_lorentz[j])

        dof_lorentzian = len(y_hb)-3

        red_lorentzian = chi_squared_lorentzian/dof_lorentzian

#####################################  LORENTZIAN RESIDUALS  #####################################

        residuals_lorentz = []

        for j in range(0,len(y_hb)):

                residuals_lorentz.append(y_hb[j] - best_lorentz[j])

        r_lor_med, r_lor_std =  np.median(residuals_lorentz), np.std(residuals_lorentz)
```

```
#################################### GAUSSIAN FITTING ####################################

        guess_gauss = [max(y_hb), 4861, 5]

        popt_g, pcov_g = curve_fit(gaussian, x_hb, y_hb, p0=guess_gauss, maxfev = 10000)

        best_gauss = gaussian(x_hb, popt_g[0], popt_g[1], popt_g[2])

#################################### GAUSSIAN CHI SQUARE ####################################

        chi_squared_gaussian = 0

        for j in range (0,len(y_hb)):

                chi_squared_gaussian = chi_squared_gaussian + np.fabs((y_hb[j] - best_gauss[j])**2/best_gauss[j])

        dof_gaussian = len(y_hb)-3

        red_gaussian = chi_squared_gaussian/dof_gaussian

#################################### GAUSSIAN RESIDUALS ####################################

        residuals_gauss = []

        for j in range(0,len(y_hb)):

                residuals_gauss.append(y_hb[j] - best_gauss[j])

        r_gau_med, r_gau_std = np.median(residuals_gauss), np.std(residuals_gauss)

#################################### PLOT OF DATA AND FITS ####################################

        plt.plot(x_hb, y_hb,linewidth=1.5, color='black', label=filename)

        plt.plot(x_hb, best_lorentz, color='red', label='Lorentzian')

        plt.plot(x_hb, best_gauss, color='blue', label='Gaussian')

        plt.legend(loc=2)

        plt.xlabel('Wavelength in Angstrom')

        plt.ylabel('Flux')

        plt.savefig(str(file1)+'.fit.pdf')
```

```
        #plt.show()

#################################### WRITING RESULTS IN OUTPUT FILE ####################################

        parameters = open('output_parameters', 'a')        #'a' appends to the already existing output file (or creates it if not present)

        #parameters = open('output_parameters', 'w')        #'w' replaces the previously existing output file (or creates it if not present)

#pameters list:

#filename, z, rms_continuum, S/N, L, lorentz_, lorentz_wavelength, lorentz_core, lorentz_chi, lorentz_redchi, lorentz_residuals_median, lorentz_residuals_stddev, G, gauss_flux, gauss_wavelength, gauss_sigma, gauss_chi, gauss_redchi, gauss_residuals_median, gauss_residuals_stddev

        parameters.write(str(filename)+'\t'+str(z[i])+'\t'+str(round(rms_cont,3))+'\t'+str(round(snratio,3))+'\t
L
\t'+str(round(popt_l[0],3))+'\t'+str(round(popt_l[1],3))+'\t'+str(round(popt_l[2],3))+'\t'+str(round(chi_squared
_lorentzian,3))+'\t'+str(round(red_lorentzian,3))+'\t'+str(round(r_lor_med,3))+'\t'+str(round(r_lor_std,3))+'\t
G
\t'+str(round(popt_g[0],3))+'\t'+str(round(popt_g[1],3))+'\t'+str(round(popt_g[2],3))+'\t'+str(round(chi_squar
ed_gaussian,3))+'\t'+str(round(red_gaussian,3))+'\t'+str(round(r_gau_med,3))+'\t'+str(round(r_gau_std,3))
+'\n')

        parameters.close()
```

Appendix – Marco Berton's luminosity code

```
import numpy as np
from scipy.stats import ks_2samp
from scipy.stats import anderson_ksamp
import matplotlib.pyplot as plt
import cosmolopy.distance as cd

cosmo = {'omega_M_0' : 0.3, 'omega_lambda_0' : 0.7, 'omega_k_0' : 0.,  'h' : 0.7}
#cosmo = cd.set_omega_k_0(cosmo)

file1 = 'gaussian.txt'
file2 = 'lorentzian.txt'

z_g = np.genfromtxt(file1, usecols=(1))
z_l = np.genfromtxt(file2, usecols=(1))

print 'z p-value:', ks_2samp(z_g, z_l)[1]

R4570_g = np.genfromtxt(file1, usecols=(2))
R4570_l = np.genfromtxt(file2, usecols=(2))

print 'R4570 p-value k-s:', ks_2samp(R4570_g, R4570_l)[1]
print 'R4570 p-value a-d:', anderson_ksamp([R4570_g, R4570_l])
print 'R4570 Gaussian mean', np.mean(R4570_g), 'median', np.median(R4570_g), '+/-', np.std(R4570_g)
print 'R4570 Lorentzi mean', np.mean(R4570_l), 'median', np.median(R4570_l), '+/-', np.std(R4570_l)
```

```
rms_cont_g = np.genfromtxt(file1, usecols=(3))

rms_cont_l = np.genfromtxt(file2, usecols=(3))


sn_g = np.genfromtxt(file1, usecols=(4))

sn_l = np.genfromtxt(file2, usecols=(4))


Mbh_g = np.log10(np.genfromtxt(file1, usecols=(8)))

Mbh_l = np.log10(np.genfromtxt(file2, usecols=(8)))


print 'MBH p-value k-s:', ks_2samp(Mbh_g, Mbh_l)[1]

print 'MBH Gaussian mean', np.mean(Mbh_g), 'median', np.median(Mbh_g), '+/-', np.std(Mbh_g)

print 'MBH Lorentzi mean', np.mean(Mbh_l), 'median', np.median(Mbh_l), '+/-', np.std(Mbh_l)


edd_g = np.genfromtxt(file1, usecols=(10))

edd_l = np.genfromtxt(file2, usecols=(10))


print 'Edd p-value k-s:', ks_2samp(edd_g, edd_l)[1]

print 'Edd Gaussian mean', np.mean(edd_g), 'median', np.median(edd_g), '+/-', np.std(edd_g)

print 'Edd Lorentzi mean', np.mean(edd_l), 'median', np.median(edd_l), '+/-', np.std(edd_l)


fo3_g = 1e-17*np.genfromtxt(file1, usecols=(12))

fo3_l = 1e-17*np.genfromtxt(file2, usecols=(12))


Lo3_g, Lo3_l = [], []


for i in range(0,len(fo3_g)):
```

```
        dl = cd.luminosity_distance(z_g[i], **cosmo)*3.08567758128e24
        Lo3_g.append(np.log10(4*np.pi*dl**2*fo3_g[i]))

for i in range(0,len(fo3_l)):
        dl = cd.luminosity_distance(z_l[i], **cosmo)*3.08567758128e24
        Lo3_l.append(np.log10(4*np.pi*dl**2*fo3_l[i]))

print 'LO3 p-value k-s:', ks_2samp(Lo3_g, Lo3_l)[1]
print 'LO3 Gaussian mean', np.mean(Lo3_g), 'median', np.median(Lo3_g), '+/-', np.std(Lo3_g)
print 'LO3 Lorentzi mean', np.mean(Lo3_l), 'median', np.median(Lo3_l), '+/-', np.std(Lo3_l)

plt.subplot(2, 1, 1)
bins = np.arange(0, 3, 0.1)
plt.hist(R4570_g, bins, color='red', label='G')
plt.axis([0, 2.5, 0, 150])
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.subplot(2, 1, 2)
plt.hist(R4570_l, bins, color='blue', label='L')
plt.axis([0, 2.5, 0, 350])
plt.xlabel('R4570', fontsize=13)
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.savefig('histo_fe.eps')
plt.show()

plt.subplot(2, 1, 1)
```

```
bins = np.arange(6, 8, 0.1)
plt.hist(Mbh_g, bins, color='red', label='G')
plt.axis([6, 8, 0, 200])
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.subplot(2, 1, 2)
plt.hist(Mbh_l, bins, color='blue', label='L')
plt.axis([6, 8, 0, 350])
plt.xlabel('$\log$M$_{BH}$ (M$_\odot$)', fontsize=13)
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.savefig('histo_bh.eps')
plt.show()

plt.subplot(2, 1, 1)
bins = np.arange(0, 1, 0.05)
plt.hist(edd_g, bins, color='red', label='G')
plt.axis([0, 1, 0, 150])
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.subplot(2, 1, 2)
plt.hist(edd_l, bins, color='blue', label='L')
plt.axis([0, 1, 0, 300])
plt.xlabel('Eddington ratio', fontsize=13)
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.savefig('histo_edd.eps')
```

```
plt.show()


plt.subplot(2, 1, 1)
bins = np.arange(39.5, 43.5, 0.25)
plt.hist(Lo3_g, bins, color='red', label='G')
plt.axis([39.5, 43.5, 0, 250])
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.subplot(2, 1, 2)
plt.hist(Lo3_l, bins, color='blue', label='L')
plt.axis([39.5, 43.5, 0, 600])
plt.xlabel('$\log$L$_{[O III]}$ (erg s$^{-1}$)', fontsize=13)
plt.ylabel('Number of sources', fontsize=10)
plt.legend(loc=1)
plt.savefig('histo_o3.eps')
plt.show()
```